# the CATO JOURNAL

AN INTERDISCIPLINARY JOURNAL OF PUBLIC POLICY ANALYSIS



| | |
|---|---|
| John C. Weicher | Mismeasuring Poverty and Progress |
| Richard B. McKenzie and Stephen D. Smith | Loss of Textile and Apparel Jobs: Is Protectionism Warranted? |
| Robert Higgs and Charlotte Twight | National Emergency and the Erosion of Private Property Rights |
| Martin T. Katzman | Environmental Risk Management through Insurance |
| Bruce L. Benson, M. L. Greenhut, and Randall G. Holcombe | Interest Groups and the Antitrust Paradox |
| Steven G. Craig and Joel W. Sailors | Interstate Trade Barriers and the Constitution |
| Peter Trapp | West Germany's Economic Policy: What Direction? |
| Charles G. Hanson | Economic Significance of British Labor Law Reform |
| Nelson Reid and Gary Lowe | Ethnic Utopianism and Market Reality in South Africa |
| Joel D. Scheraga | Establishing Property Rights in Outer Space |
| Gary M. Anderson and Dolores T. Martin | The Public Domain and Nineteenth Century Transfer Policy |
| William F. Shughart II | Don't Revise the Clayton Act, Scrap It! |
| Charles W. Baird | Labor Law and Labor-Management Cooperation: Two Incompatible Views |



VOLUME 6 NUMBER 3 • PUBLISHED BY THE CATO INSTITUTE • WINTER 1987



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# the CATO JOURNAL

VOLUME 6 NUMBER 3 • PUBLISHED BY THE CATO INSTITUTE • WINTER 1987

Editor
JAMES A. DORN

Assistant Editor
CATHERINE ENGLAND

Production
DAVID LAMPO

EDITORIAL BOARD

TERRY L. ANDERSON
Montana State University

PETER H. ARANSON
Emory University

CHARLES W. BAIRD
California State University
at Hayward

T. E. BORCHERDING
Claremont Graduate School

YALE BROZEN
University of Chicago

KARL BRUNNER
University of Rochester

JAMES M. BUCHANAN
Center for Study of
Public Choice
George Mason University

ARTHUR A. EKIRCH
State University
of New York at Albany

RICHARD A. EPSTEIN
University of Chicago

STEVE H. HANKE
Johns Hopkins University

C. LOWELL HARRISS
Columbia University

FRIEDRICH A. HAYEK
University of Freiburg

JOSEPH P. KALT
Harvard University

LEONARD P. LIGGIO
Institute for
Humane Studies

WILLIAM A. NISKANEN
Cato Institute

SAM PELTZMAN
University of Chicago

ALVIN RABUSHKA
Hoover Institution
Stanford University

EARL C. RAVENAL
Georgetown University

VERNON L. SMITH
University of Arizona

ROLAND VAUBEL
Mannheim University

KAREN I. VAUGHN
George Mason University

RICHARD E. WAGNER
Florida State University

LAWRENCE H. WHITE
New York University

LELAND B. YEAGER
Auburn University

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The CATO JOURNAL (ISSN 0273-3072) is published in the spring/summer, fall, and winter by the Cato Institute, 224 Second Street SE, Washington, D.C. 20003. Annual subscriptions are $18.00 for both individuals and institutions. Single copies are $6.00. Two-year subscriptions are $36.00 and three-year subscriptions are $54.00. Foreign subscribers should add $5.00 per year for regular delivery and $10.00 per year for airmail delivery.

Correspondence regarding subscriptions, changes of address, procuring back issues, advertising and marketing matters, etc., should be addressed to the Publications Department. All other correspondence, including requests to quote or reproduce material, should be addressed to the editor.

Unsolicited manuscripts cannot be returned and will be acknowledged only if accompanied by a stamped, self-addressed envelope. There is no submission fee. Articles should focus on specific public policy issues. Accepted manuscripts must conform to the CATO JOURNAL's style requirements. (A style sheet is available on request.)

Authors should submit three copies of their papers or communications to the editor. An abstract of no more than 250 words should also be included to aid referees. Manuscripts should have wide margins and be double-spaced throughout, including footnotes.

Unsolicited book reviews cannot be returned and will be acknowledged only if they are accepted for publication. They should conform to the above requirements, and should provide an in-depth analysis of major studies of public policy. Reviews should range from 800 to 1,000 words.

The views expressed by the authors of the articles are their own and are not attributable to the editor, the editorial board, or the Cato Institute.

The CATO JOURNAL is indexed and abstracted in the *Journal of Economic Literature, Current Contents/Social and Behavioral Sciences, Social Sciences Citation Index,* and the ABI/INFORM database. The CATO JOURNAL is available on microfilm from University Microfilms International, Ann Arbor, Michigan 48106.

Printed in the United States of America.

Copyright © 1987 by the Cato Institute.
All rights reserved.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

**John C. Weicher**
Mismeasuring Poverty and Progress 715

**Richard B. McKenzie and Stephen D. Smith**
Loss of Textile and Apparel Jobs: Is Protectionism Warranted? 731

**Robert Higgs and Charlotte Twight**
National Emergency and the Erosion of Private Property Rights 747

**Martin T. Katzman**
Environmental Risk Management through Insurance 775

**Bruce L. Benson, M. L. Greenhut, and Randall G. Holcombe**
Interest Groups and the Antitrust Paradox 801

**Steven G. Craig and Joel W. Sailors**
Interstate Trade Barriers and the Constitution 819

**Peter Trapp**
West Germany's Economic Policy: What Direction? 837

**Charles G. Hanson**
Economic Significance of British Labor Law Reform 851

**Nelson Reid and Gary Lowe**
Ethnic Utopianism and Market Reality in South Africa 869

**Joel D. Scheraga**
Establishing Property Rights in Outer Space 889

**Gary M. Anderson and Dolores T. Martin**
The Public Domain and Nineteenth Century Transfer Policy 905

**William F. Shughart II**
Don't Revise the Clayton Act, Scrap It! 925

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

**Charles W. Baird**
Labor Law and Labor-Management Cooperation:
Two Incompatible Views 933

**BOOK REVIEWS**

*Perspectives on Safe and Sound Banking: Past, Present and Future*
George J. Benston, Robert A. Eisenbeis, Paul M. Horvitz, Edward J. Kane, and George G. Kaufman
*Reviewed by* **Gillian Garcia and Kim Staking** 951

*Superfairness*
William J. Baumol
*Reviewed by* **Philip Porter** 954

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Mismeasuring Poverty and Progress

*John C. Weicher*

## The Mistake

In 1969 the federal government decided to change the method of adjusting the poverty line income each year. Instead of changing in proportion to the change in the cost of food, the poverty line henceforth was to be adjusted by the change in the overall Consumer Price Index. And so it has been. This technical change was a mistake, and the mistake has colored and distorted perceptions of the poverty problem.

The change was a mistake because the CPI mismeasured the cost of owning a home—and badly mismeasured it during periods of inflation. For any given year the CPI measured only the cost of buying a home, not the cost of living in it. The cost of homeownership was calculated as the full amount that would be paid by home buyers during the period that they could be expected to own the home, that is, the purchase price and the total amount of all mortgage payments for 15 years.[1] These are large amounts, compared with the current incomes and expenditures of most households, but they were treated as expenses only for those who actually bought houses, about 3 percent of all households in a given year. For the other 60 percent that already owned a home and did not buy another one during the year, no homeownership costs were computed, other than current outlays for property taxes, insurance, and repairs.[2]

---

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

The author is F. K. Weyerhaeuser Scholar in Public Policy Research at the American Enterprise Institute. An earlier version of this paper was presented at the American Enterprise Institute Public Policy Week, December 1985. The author gratefully acknowledges the assistance of Daniel H. Weinberg of the U.S. Department of Health and Human Services in providing the data in Table 2.

[1]The cost of homeownership also includes annual expenditures for property taxes, maintenance, and insurance; these do not pose measurement problems and are omitted from this discussion.

[2]For a more extensive discussion of the problems with the homeownership component of the CPI, see Cagan and Moore (1981, especially pp. 32–43).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

This procedure gave too much weight to changes in housing prices. During periods of inflation, these prices rise more rapidly than prices in general, because houses are an asset as well as a consumer good. Inflationary pressures during the late 1960s and the 1970s, for instance, caused problems primarily for those seeking to buy a house for the first time. Families that owned their homes before the inflation developed (that is, most families) enjoyed substantial real capital gains, making it possible for them to buy bigger and better houses. For such families, the price increase reported in the CPI was an increase in wealth, not an increase in the cost of living.[3] In fact, houses proved to be the best hedge against inflation available to most households. (Even greater real price increases occurred for other real assets, such as precious metals, objets d'art, coins, and stamps.)

The CPI also overemphasized changes in mortgage interest rates, which are notoriously volatile. During the early 1970s, the effect of these changes on the CPI was not too great, because mortgage lenders, like most lenders, underestimated the strength and severity of inflation, and did not raise rates to compensate for it. By the end of the 1970s, however, lenders had learned their lesson, and mortgage rates rose dramatically (if erratically). The impact of this increase on the CPI can be judged from the following example: in 1980 the CPI increased by over 13 percent, but in the month of July it did not rise at all; a sharp but temporary drop of 5 percent in mortgage rates that month was enough to offset a double-digit increase across the rest of the marketbasket.

By the end of 1979, the homeownership component amounted to 25 percent of the entire CPI, far in excess of the share that owner-occupied housing actually represented in the typical household budget, and five times the weight of the rental component. This figure was a gross distortion of actual behavior. In 1979, according to the Annual Housing Survey, the median ratio of monthly housing costs to income was 17 percent for homeowners; since homeowners amounted to slightly less than two-thirds of all households, homeownership costs were about 11 percent for the "typical" household. The median rent-income ratio was 26 percent, or about 8 percent of income for the typical household.

The CPI's measurement problem was widely recognized. Several housing market analysts constructed alternative indices that included the increase in the value of the home as an offset to the increase in its price. These measures showed a much smaller increase in the cost of owning a home during the 1970s—in some years, a decrease—

[3]For further discussion of housing and inflation, see Weicher (1984, pp. 155–204).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and resulted in a much smaller increase in the overall CPI.[4] The Bureau of Labor Statistics was aware of the measurement problem at least by 1975 but was unable to correct it until 1983.[5] The bureau now measures the cost of buying a home as its "rental equivalence," or the amount that would have to be paid if the owner actually rented the home from someone else. The rental equivalence measure is used in the National Income and Product Accounts and has been used in the CPI for the last four years.

## Consequences of the Mistake for the CPI

The distortion of the CPI measurement of homeownership costs did not matter much in the early years, but gradually it added up to something fairly substantial. Table 1 shows the official CPI and an experimental version using the rental equivalence approach for the years 1967–85.[6] The official CPI went up faster in years of accelerating inflation, and at least as fast as the experimental index even during recessions. As late as 1978, the official index was only about 5 percentage points higher than the experimental index. But as inflation accelerated sharply at the end of the decade, the discrepancy doubled within three years. By 1981, the official CPI was 10 percentage points higher than the experimental index. The difference remained unchanged in 1982, when the process of disinflation was gathering momentum; housing prices in that year rose more slowly than other goods and services, but mortgage rates continued rising for most of the year.

The continued disinflation and the decline in both real housing prices and mortgage rates would have caused the official index to rise more slowly than the experimental index, thus bringing them closer. (The old index would have risen by about 3.0 percent in 1983 and 3.4 percent in 1984, compared with 3.2 and 4.2 percent, respec-

[4]See, for example, Dougherty and Van Order (1982, pp. 154–64).

[5]For a discussion of the problem by a BLS economist and references to much of the literature, see Greenlees (1982, pp. 1–24).

[6]The experimental index, or CPI-U-X1, for the years since 1967 was published in the 1982 and 1983 *Economic Report of the President*. The same technical problem with the homeownership index existed before 1967 as well but was insignificant for practical purposes. Between 1953 (the first year homeownership costs were measured in the CPI) and 1967, the homeownership index increased by 7.2 percent more than the rent index; the overall CPI would have been about 1 percent lower in 1967 if the experimental index had been used since 1953.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 1*

OFFICIAL AND EXPERIMENTAL CONSUMER PRICE INDEX, 1967–85

| Year | Consumer Price Index | | Ratio of Official to Experimental |
|---|---|---|---|
| | Official (CPI-U) | Experimental (CPI-U-X1) | |
| 1967 | 100.0 | 100.0 | 1.000 |
| 1968 | 104.2 | 103.7 | 1.005 |
| 1969 | 109.8 | 108.3 | 1.014 |
| 1970 | 116.3 | 113.6 | 1.024 |
| 1971 | 121.3 | 118.5 | 1.024 |
| 1972 | 125.3 | 122.1 | 1.026 |
| 1973 | 133.1 | 129.7 | 1.026 |
| 1974 | 147.7 | 142.8 | 1.030 |
| 1975 | 161.2 | 154.6 | 1.043 |
| 1976 | 170.5 | 163.5 | 1.043 |
| 1977 | 181.5 | 173.9 | 1.044 |
| 1978 | 195.4 | 185.7 | 1.052 |
| 1979 | 217.4 | 203.6 | 1.068 |
| 1980 | 246.8 | 226.4 | 1.090 |
| 1981 | 272.4 | 247.9 | 1.099 |
| 1982 | 289.1 | 263.0 | 1.099 |
| 1983 | 298.4 | 271.5 | 1.099 |
| 1984 | 311.0 | 283.0 | 1.099 |
| 1985 | 322.2 | 293.1 | 1.099 |

SOURCE: *Economic Report of the President* (1983, Tables B-54 and B-56; 1986, Table B-55).

tively, by the new method.)[7] But this correction was not made because the official index adopted the rental equivalence approach in 1983, thereby locking in the error, presumably forever, since the CPI is never changed.

## Consequences of the Mistake for the Poverty Rate

As the official CPI overstated the increase in prices for all consumers between 1967 and 1982, so the official poverty thresholds, adjusted

[7]Precise calculation of the CPI-U by the old method is not possible from published data, but a close approximation can be made. The old cost of homeownership was continued in the CPI-W in 1983 and 1984. I have combined the changes in the components of homeownership cost and other components of the shelter index (primarily rent) from the CPI-W with the published CPI-U excluding shelter, using the 1982 weights for these components from the two indices as the base period quantities. Other estimating techniques are likely to give very similar results, so far as I can judge from the published data.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

by the CPI, overstated the increase in prices for the poor. Indeed, the official CPI makes far less sense as a measure of the "cost of living" for the poor than it does for the population in general. Sixty percent of the poor are renters, and half of those who own are at least 60 years old; the cost of buying a home is surely irrelevant for most of them. Very few poor people buy a home in any year. Data from the Annual Housing Survey for 1981 suggest that less than 1.4 percent of poor households bought a home, compared with 3.2 percent of all households.[8] The CPI has been an inaccurate measure of the behavior of the poor in the housing market.

The overstatement of the official poverty thresholds means that the number of poor people and the poverty rate have also been overstated, and continue to be overstated, if one assumes that the current version of the CPI is appropriate and the one that should have been used all along.[9] Again the error is small at first, but it accumulates over time and becomes large by the end of the 1970s. For 1982, the poverty rate and the number of poor are too high by one-eighth—almost 2 percentage points and over 4 million people. This discrepancy has remained essentially unchanged and will continue as such in the future, unless the mistake in the calculation of the official poverty thresholds is corrected. Table 2 shows the yearly official and adjusted figures for both the rate and the number of poor people.[10]

[8]The AHS does not classify households by poverty status. It does have the income distribution for home buyers during the previous year, but not cross-classified by household size. The typical buying household in 1981 had 2.7 persons; the poverty income for 3-person households was just above $7,000. Fewer than 1.4 percent of buyers had incomes of less than $7,000.

[9]The rental equivalence measures differ between the experimental index and the new CPI. For the former, the rent component of the CPI was used; for the latter, a new sample of rental housing units was drawn, with similar characteristics to owner-occupied housing. Thus the new CPI is not quite the same as the experimental index, but the differences are small. For the first three years, the new homeowners' equivalent rent index rose by 13.1 percent, the rent index by 14.7 percent. Had the latter been used in the CPI instead of the former—that is, had the experimental CPI become the new official index—the CPI would have been two-tenths of 1 percentage point higher in 1985. This suggests that the new official index would have been close to the experimental index between 1967 and 1982, though it is not proof.

[10]The figures shown in Table 2 differ slightly from those in the previous version of this paper presented at the seminar on "The New Meaning of Poverty," American Enterprise Institute Public Policy Week, December 1985. The adjusted rates were provided by Daniel H. Weinberg of the U.S. Department of Health and Human Services, and calculated by him from the data tapes for the Current Population Survey. I have calculated the number of poor people from Weinberg's rates. My earlier figures were calculated from the published Consumer Income annual reports, Series P-60 of the *Current Population Reports*. Because the published data are grouped by income class,



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 2*

OFFICIAL AND ADJUSTED POVERTY, 1965–85

| Year[a] | Poverty Rate (Percent) | | Number of Poor People (Millions) | |
|---|---|---|---|---|
| | Official | Adjusted | Official | Adjusted |
| 1965 | 17.3 | 17.3 | 33.2 | 33.2 |
| 1966 | 14.7 | 14.7 | 28.5 | 28.5 |
| 1967 | 14.3 | 14.3 | 27.8 | 27.8 |
| 1968 | 12.8 | 12.7 | 25.4 | 25.2 |
| 1969 | 12.2 | 11.9 | 24.1 | 23.5 |
| 1970 | 12.6 | 12.2 | 25.4 | 24.6 |
| 1971 | 12.5 | 12.1 | 25.6 | 24.8 |
| 1972 | 11.9 | 11.4 | 24.5 | 23.5 |
| 1973 | 11.1 | 10.6 | 23.0 | 22.0 |
| 1974 | 11.6 | 11.0 | 23.4 | 22.3 |
| 1975 | 12.3 | 11.5 | 25.9 | 24.3 |
| 1976 | 11.8 | 11.0 | 25.0 | 23.4 |
| 1977 | 11.6 | 10.7 | 24.7 | 22.9 |
| 1978 | 11.4 | 10.4 | 24.5 | 22.4 |
| 1979 | 11.6 | 10.5 | 26.1 | 23.4 |
| 1980 | 13.0 | 11.5 | 29.3 | 25.9 |
| 1981 | 14.0 | 12.2 | 31.8 | 27.7 |
| 1982 | 15.0 | 13.2 | 34.4 | 30.3 |
| 1983 | 15.3 | 13.5 | 35.5 | 31.2 |
| 1984 | 14.4 | 12.7 | 33.7 | 29.7 |
| 1985 | 14.0 | 12.3 | 33.0 | 29.1 |

[a]No adjustment for 1965 through 1967.
SOURCE: U.S. Department of Health and Human Services; see supra, n. 10.

To the poor themselves, this is a distinction without a difference. The poverty line is not like the Iron Curtain—if you are close to it, it matters little which side you are on. It is a reasonable effort to estimate the cost of a minimally adequate standard of living, and there is some logic to it, but a family that is $10 above the poverty line is not qualitatively better off than one $10 below it. The latter will not starve; the former will not maintain itself in blooming health. At the same time, a family that is $5,000 below the line is a lot worse off than one just below it, but both are counted equally.

some rounding error was introduced; however, Weinberg's figures differ from my earlier ones by more than 0.1 in only two of the 19 years, and the pattern in the table is unchanged.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Moreover, even with the adjustment, there is still more than enough poverty. Having almost one-eighth of the population in poverty in 1985 (the adjusted rate) is better than having almost one-seventh (the official rate), but it certainly leaves room for improvement.

## Consequences of the Mistake for "the Poverty Problem"

If the overstatement of poverty does not make much difference quantitatively, it does matter politically. It has affected our view of the problem of poverty, of what policies should be adopted to ameliorate it, and even our view of our society, its achievements and failures. Despite its recognized limitations, the poverty rate has become *the* newsworthy number about the poor.[11] In that sense, it is like the unemployment rate and the rate of housing starts: such measures need to be qualified and supplemented, and sometimes they are not the best guides to the situation in the society or the economy.

Politically, the important difference between the official and the adjusted poverty rates is not the level but the pattern of change. The adjustment modifies the pattern in two important ways: the long-term trend and the experience of the 1980s both look different. The official rate reached its lowest level in 1973; it rose during the 1973–75 recession but did not return to the 1973 level during the subsequent recovery. The lowest adjusted rate, however, was reached in 1978 and nearly maintained in 1979. In both years, the rate was below the 1973 rate. In the course of the business cycle, the poverty rate had been reduced. Both the official and the adjusted rates rose sharply in the early 1980s, but the official rate rose more, even though the two CPIs rose together after 1981. From its cyclical low in 1978, the official rate rose almost 4 percentage points to 1983, and is still about 2.5 percentage points above the low point. Over the same period, the adjusted rate rose about 3 percentage points and then came down to within 2 percentage points of its lowest level.

These differences are important because there is a widespread popular perception, shared by people with quite different ideologies, that the United States stopped making progress against poverty in the 1970s. My sense is that the perception is ultimately rooted in the fact that the official poverty rate has not fallen for more than a decade.

Consider, for example, Charles Murray's well-known critique of welfare programs, *Losing Ground*. Poverty is discussed in the first

[11]This point was made by each of the papers delivered at the Census Bureau's Conference on the Measurement of Noncash Benefits, Williamsburg, December 1985. See, for example, Chiswick (1985, p. 37).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of his chapters describing the deteriorating condition of the poor and disadvantaged after the creation of the Great Society programs. Murray (1984, p. 58) notes the decline in the official poverty rate during the 1960s and goes on to say:

> Then, after two decades of reasonably steady progress, improvement slowed in the late sixties and stopped altogether in the seventies. The proportion dipped to its low point, 11 percent, in 1973. A higher proportion of the American population was officially poor in 1980 than at any time since 1967.

If the poverty rate had been measured then as it is now, Murray's last two sentences would have to read differently: "The proportion dipped to its low point, 10.5 percent, in 1979. A higher proportion of the American population was officially poor in 1980 than at any time since 1975." That does not sound so bad. Indeed, with those numbers, the sentence would certainly have been written differently.

Consider, alternatively, Lester Thurow's recent proposal for structural reform of the American economy, *The Zero-Sum Solution*. In describing the changing position of the middle class, he states (1985, p. 61):

> Many fall out of the middle class with enough momentum to land in poverty. The percentage of the population living below the poverty line reached a low of 11.1 percent in 1973, hovered between 11 and 12 percent until 1979, and then began a steep ascent—rising to 15.2 percent in 1983. More than 35 million Americans are now living in poverty. One has to go back to 1965 to find a year with a higher incidence of poverty.

Again the paragraph would have to be rewritten. It would still be true that the 1983 poverty rate was high compared with recent experience, but the problem would look more recent and be more clearly related to the recessions of the early 1980s.

The parallels between Murray and Thurow go further. In each case the adjustment in the poverty rate would not affect their analyses of other—in my view, more important—aspects of their arguments. Murray goes on to talk about crime, education, and demographic changes; Thurow discusses increases in economic inequality (not the same as increases in poverty). Nevertheless, both try to buttress their arguments by an appeal to the stubbornly high and rising poverty rate.

Consider finally what the newspaper headlines about each year's poverty statistics might have been in the late 1970s. For 1977, perhaps, "Poverty Rate Falls Slightly, Nears All-Time Low"; for 1978, more certainly, "Poverty Rate Is Lowest Ever"; for 1979, "Poverty RateUnchanged, Remains at Record Low." Such headlines would



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

not create or reinforce negative perceptions of the problem. We might still be concerned that we were not making more rapid progress, but that is different from believing that we are not making any progress.

In fact, we may still be making progress in the 1980s. The lowest poverty rate of this decade occurred at the peak of the last business cycle; since then, we have experienced two recessions, the second of them quite severe, with a very mild and short-lived recovery in between. The recovery from the second recession began at the end of 1982. The current economic recovery has not yet run its course. The consensus of economic forecasters now seems to be that the next recession will not begin before 1988, meaning that 1987 at the earliest will be the next cyclical peak year. Thus it is possible that the poverty rate at the next peak will be lower than it was in 1979. It has, of course, risen very strongly since then, more so than in any previous recession, but the last recession was also the worst of the postwar period.

Ironically, if the CPI had not been changed in 1983, the poverty rate would be lower now, because of the decline in real housing prices and in mortgage rates. Measured with the old homeownership cost index, the poverty rate would have been about 15.2 percent in 1983, instead of 15.3 percent; for 1984, it would have come down to about 14.2 percent, instead of 14.4 percent.[12] The old index was discontinued in mid-1985, but it is clear that the CPI and the poverty rate would have been markedly lower for 1985 if it had still been in use.[13]

Continuing to measure poverty by the old index certainly would not have been right, but it would have partly offset the accumulated discrepancy and it would have looked better. Even the process of correcting the original mistake has contributed to the impression that poverty is an intractable problem.

The significance of the adjustment should not be overemphasized. It is still true that progress against poverty was much slower in the 1970s than in the preceding two decades, as Murray pointed out.

[12]These figures are derived from the estimated change in the old official index (see supra, n. 7).

[13]For five of the first six months of 1985, the new homeowners' equivalent rent index rose more rapidly than the old homeownership index, on a year-over-year basis. In the second half of the year, the rental equivalence measure rose by 3.0 percent. Over the same period, mortgage interest rates *declined* by about 7 percent, and housing prices increased by at most 2.1 percent. These are the two largest components of the old homeownership index. (The old BLS housing price index was discontinued after June 1985; the Census Bureau's "New One-Family House Price Index" increased by 2.1 percent in the last half of the year. Historically, the Census index has increased more rapidly than the BLS index.)



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Several explanations have been offered, such as the increasing proportion of female-headed households and the failure of means-tested cash benefits to keep pace with inflation (the latter is still true for the experimental CPI). These are certainly valid and they help explain the changes in the poverty rate over time, but they contribute less to an explanation of changes in the adjusted poverty rate. By the adjusted rate, poverty looks a little more like a cyclical or economic problem, a little less like a secular or structural one.[14]

## Who Are the Poor (Adjusted)?

The correction in the poverty rate would not greatly affect the composition of the poverty population, but it would accentuate the trends many analysts have recognized. Table 3 compares the changes in the characteristics of the poor between 1967 and 1984 (the latest year available at this writing), using the official and the adjusted poverty thresholds. The most obvious trend, according to both measures, is the much-discussed "feminization of poverty," in which an increasing share of the poor are members of female-headed families. Children are a slightly declining fraction of the poor, but this is the net effect of disparate trends for families headed by males (nearly all of them married couples) and families headed by women. There has been a sharp decline in poverty among the elderly, and an increase for members of "other races." All these trends are a little more pronounced when the poverty population is calculated on the basis of the adjusted poverty thresholds.

It may seem paradoxical that the adjusted poverty rate shows a more cyclical pattern, while the adjusted poverty population has a higher concentration of individuals likely to be out of the labor force. The data, however, are internally consistent. People who are close to the official poverty line tend to be those who are more willing to work and better able to find a job when the economy is strong. More of them, but not all, are therefore above the adjusted poverty line. Women without husbands, especially such women with children, are typically further below the poverty line to begin with, and are less

[14]Although the foregoing discussion refers specifically only to the official poverty rate, it would apply also to other concepts, such as pretransfer or latent poverty, or the poverty rate after in-kind benefits are included. Published data do not permit easy adjustment of these other concepts, but it seems likely that pretransfer poverty would have declined slightly during the 1970s (though not as fast as in the 1960s and earlier), instead of rising; and poverty net of in-kind benefits would have fallen more sharply than the various published estimates report. The Census data available since 1979 show all measures of poverty net of in-kind benefits moving very closely with the official rate; perhaps the same would be true for adjusted poverty rates.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 3*

COMPOSITION OF THE POVERTY POPULATION, 1967–84
(PROPORTION OF THE POOR IN EACH CATEGORY)

| Characteristics | Official | | Adjusted | |
|---|---|---|---|---|
| | 1967 | 1984 | 1967 | 1984 |
| *Families:* | | | | |
| Female-Headed | 31.3 | 48.1 | 31.3 | 49.8 |
| *Individuals:* | | | | |
| In Families | 82.0 | 78.5 | 82.0 | 78.9 |
| Male-Headed | 57.2 | 43.4 | 57.2 | 42.3 |
| Female-Headed | 24.8 | 35.1 | 24.8 | 36.6 |
| In Unrelated Subfamilies | NA[a] | 1.9 | NA[a] | 2.0 |
| Unrelated Men | 4.7 | 7.6 | 4.7 | 7.7 |
| Unrelated Women | 13.3 | 12.0 | 13.3 | 11.4 |
| *Children* | 41.2 | 38.4 | 41.2 | 39.2 |
| In Male-Headed Families | 25.9 | 18.5 | 25.9 | 18.0 |
| In Female-Headed Families | 15.3 | 20.1 | 15.3 | 21.2 |
| *The Elderly* | 15.5[b] | 9.9 | 15.5[b] | 8.7 |
| *Race* | | | | |
| White | 68.4 | 68.1 | 68.4 | 67.3 |
| Black | 30.5 | 28.2 | 30.5 | 28.9 |

[a]Data not available separately for 1967; included in category "In Families."
[b]Underestimate; includes only household heads and unrelated individuals, omitting elderly spouses, elderly brothers or sisters of the head or spouse, and other elderly relatives.
SOURCE: U.S. Bureau of the Census, "Characteristics of the Population below the Poverty Level" (1977, no. 119, Table 7 [for 1967]; 1984, no. 152, Table 6).

likely, or at least slower, to get jobs during an economic upturn. If these trends in the poverty population continue, eventually the poor may be simply a "hard core" outside the labor force, but that is not the situation now, even by the adjusted rate.

The elderly are an exception. They are concentrated between about 75 percent and 150 percent of the official poverty thresholds, so the adjustment removes more of them from the poverty population, but their situation is not particularly sensitive to the economy.

## Where Are the Poor (Adjusted)?

The error may distort the geography of poverty. Published Census reports show that the number of poor people living in big cities rose



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

between 1970 and 1980, and their poverty areas expanded, covering more territory and including more people. The apparent increase in the number of poor people probably results in large part from the measurement problem. Using the official poverty thresholds, the number of poor people in big cities rose from 6.0 to 6.7 million; using the adjusted thresholds, it rose from 5.7 to 5.8 million. Similar calculations cannot be performed for poverty neighborhoods from published data, but it is likely that a number of Census tracts that just barely qualify for poverty status according to the official measure would no longer be poverty areas. A poverty area consists of one or more Census tracts in which at least 20 percent of the population is poor; if the measured number of poor people in central cities were adjusted downwards by over 10 percent, then the number of Census tracts classified as poverty areas would surely be smaller, and the territorial spread of poverty would look less foreboding on municipal maps. Again, this correction would not change the well-being of any individual, but would affect public perceptions about the changing patterns of poverty.[15]

At the same time, any adjustment would surely not alter the perception of who lives in poverty areas. The official statistics show an increasing concentration of the poor, particularly the black and Hispanic poor, within big-city poverty neighborhoods. That is still true. Moreover, the adjustment would probably result in a substantial decline in the measured concentration of the white poor, while still showing a substantial increase among blacks and Hispanics. Perceptions of the concentration of poverty in urban ghettos would probably be sharpened.

## Mismeasuring Family Incomes: The Slipping Middle Class?

The distortion in the CPI affects broader measures of economic well-being as well as the poverty rate. The CPI is used to construct the standard measures of real family and household incomes. Because the adjusted CPI rose more slowly than the official index, the growth in real income has been understated. Measured consistently according to the current CPI methodology, real incomes from 1965 to 1981 were lower than they appear to be according to the current CPI.

[15]For an analysis of urban poverty based on the officially measured growth of poverty within big cities and poverty neighborhoods, see Richard P. Nathan, "The Underclass—Will It Always Be With Us?" paper prepared for a symposium at the New School for Social Research, 14 November 1986.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Table 4 shows real median family income, the most commonly cited income figure, adjusted by the official and the corrected CPI. The patterns are similar to those for the poverty rate, and so are the consequences of the correction. Using the official CPI, 1973 is the highest income year; using the corrected index, 1979 is the highest. As with the poor, it appears that the middle class has been losing ground in economic terms for more than a decade. This has occasioned some academic analysis and much popular discussion about the "shrinking middle class."[16]

*TABLE 4*

REAL MEDIAN FAMILY INCOME, 1967–85
(1985 DOLLARS)

| Year | Official (CPI-U) | Adjusted[a] (CPI-U-X1) |
|---|---|---|
| 1967 | $25,560 | $23,252 |
| 1968 | 26,691 | 24,398 |
| 1969 | 27,681 | 25,529 |
| 1970 | 27,339 | 25,458 |
| 1971 | 27,319 | 25,439 |
| 1972 | 28,584 | 26,684 |
| 1973 | 29,172 | 27,233 |
| 1974 | 28,145 | 26,482 |
| 1975 | 27,431 | 26,009 |
| 1976 | 28,267 | 26,815 |
| 1977 | 28,419 | 26,982 |
| 1978 | 29,087 | 27,842 |
| 1979 | 29,029 | 28,197 |
| 1980 | 27,446 | 27,217 |
| 1981 | 26,481 | 26,470 |
| 1982 | 26,116 | 26,116 |
| 1983 | 26,642 | 26,642 |
| 1984 | 27,376 | 27,376 |
| 1985 | 27,735 | 27,735 |
| Rates of Increase (%) | | |
| 1967–85 | +8.5 | +19.3 |
| 1973–85 | −4.9 | + 1.8 |
| 1979–85 | −4.5 | − 1.6 |

[a]New official CPI-U used after 1982.
SOURCES: *Economic Report of the President* (1983, Tables B-54 and B-56; 1986, Table B-55); U.S. Bureau of the Census, "Characteristics of the Population below the Poverty Level" (1977, no. 119, Table 7 [for 1967]; 1984, no. 152, Table 6).

[16]See, for example, "Is the Middle Class Shrinking?" *Time* (3 November 1986): 54–56;



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The official data also give a stronger impression of continuing secular stagnation. It will take perhaps three more years' income growth for the official median income figure to reach the 1973 peak; it should take at most two years, probably only one, to exceed the corrected 1979 high point. The current economic recovery would have to last unusually long for the former to happen, whereas it may well have already lasted long enough for the latter. As with the poverty rate, the corrected pattern of real median family income looks more sensitive to the business cycle.[17]

Again, the correction should be kept in perspective. Real income increased much more rapidly before 1973 than it did afterward, by any measure; the change in the trend has been striking. There is, however, a political difference between slightly declining and slightly rising real incomes.

## Policy Implications

The policy implications are in the eye of the beholder. Consider the debate over President Reagan's changes in means-tested programs—the "fairness issue"—in the 1984 campaign. Both sides could find ammunition. The President and his supporters could have noted that benefits were rising more rapidly than the poverty line, and they should be trimmed back. This argument is most relevant to food stamps, housing subsidies, and Medicaid, where Reagan proposed the largest changes. His critics could have rallied behind Bert Lance's adage, "If it ain't broke, don't fix it," and argued that the means-tested benefit programs were not broke and did not need to be changed, because the poverty rate was still coming down.

My guess is that benefit levels in the means-tested programs would probably not be very different. To begin with, they are not tied to the poverty line. Eligibility for food stamps and school lunches has been stated as a fraction of the poverty line, so that fewer people would have been eligible. But when the President requested a lower eligibility limit, Congress approved it for one program and not the

---

and Bradbury (1986, pp. 41–55). Bradbury's article presents especially detailed calculations, using the official CPI.

[17]The difference in the cost of homeownership between the old CPI and the GNP deflator also contributes to differences in income measures calculated from different data series. An example is presented in Ryscavage (1986, pp. 24–29). Ryscavage focuses on differences between real family income, reported by the Census Bureau, and real per capita income, reported by the Bureau of Economic Analysis; the latter is deflated by the GNP deflator (which is close to the present CPI) rather than the CPI. He attributes about a quarter of the difference in growth rates to the difference in deflators.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

other. There was no consensus that too many people were automatically entitled to benefits.

The biggest difference would be in Social Security, because it has been indexed to the official CPI since 1974. Year by year since then, benefits would have increased more slowly. By now they would be almost 7 percent lower. That would mean $12 billion less in outlays this year, a tidy sum that would contribute its mite to reducing the federal deficit, and a cumulative $100 billion over the last decade, reducing the national debt by about 4 percent.

All this is conjectural. Congress has been solicitous of the elderly, and might have increased benefits by statute if the current CPI had been in use. Means-tested benefit levels probably would not be much different than they are; the President would probably have proposed the same structural changes, since many of them were based on his experience as governor of California before inflation became a serious problem. Policies might be essentially the same.

I think, however, that our perception of the poverty problem would be different. We would feel a little better about our progress, our economy, and our society. There might be concrete differences, but they would matter less than these intangibles.

## References

Bradbury, Katharine L. "The Shrinking Middle Class." *New England Economic Review* (September/October 1986): 41–55.

Cagan, Phillip, and Moore, Geoffrey H. *The Consumer Price Index: Issues and Alternatives.* Washington, D.C.: American Enterprise Institute, 1981.

Chiswick, Barry R. "Evaluation of Census Bureau Procedures for the Measurement of Noncash Benefits and the Incidence of Poverty." In *Proceedings of the Conference on the Measurement of Noncash Benefits, vol. 1.* Washington, D.C.: Government Printing Office, 1985.

Dougherty, Anne, and Van Order, Robert. "Inflation, Housing Costs, and the Consumer Price Index." *American Economic Review* 72 (March 1982): 154–64.

*Economic Report of the President* (for 1983 and 1986). Washington, D.C.: Government Printing Office.

Greenlees, John S. "An Empirical Evaluation of the CPI Home Purchase Index, 1973–1978." *AREUEA Journal* 10 (Spring 1982): 1–24.

"Is the Middle Class Shrinking?" *Time* (3 November 1986): 54–56.

Murray, Charles. *Losing Ground.* New York: Basic Books, 1984.

Nathan, Richard P. "The Underclass—Will It Always Be With Us?" Paper prepared for sysmposium at the New School for Social Research, 14 November 1986.

Ryscavage, Paul. "Reconciling Divergent Trends in Real Income." *Monthly Labor Review* (July 1986): 24–29.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Thurow, Lester C. *The Zero-Sum Solution: Building a World-Class Economy*. New York: Simon and Schuster, 1985.

U.S. Bureau of the Census. "Characteristics of the Population below the Poverty Level." *Current Population Reports*, Series P-60. Washington, D.C., 1977 and 1984.

Weicher, John C. "Disinflation in the Housing Market." In *Essays in Contemporary Economic Problems: Disinflation*. Edited by William Fellner. Washington, D.C.: American Enterprise Institute, 1984.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Loss of Textile and Apparel Jobs: Is Protectionism Warranted?

*Richard B. McKenzie and Stephen D. Smith*

## Introduction

Backers of the recently vetoed Textile and Apparel Trade Enforcement Act of 1985 (S.680) sought to reduce textile imports by as much as 36 percent and apparel imports by as much as 20 percent from their 1985 levels.[1] If it had been enacted, the bill would have, according to one estimate, raised wholesale prices for textile imports by 33 percent and for apparel imports by 16 percent. It would have also added at least $3.4 billion to the annual consumer cost of textile protectionism already estimated at more than $20 billion a year (see Megna and Emrich 1985).

These proposed textile and apparel trade restrictions have been tendered on the proposition that expanding textile and apparel imports have caused the closure of as many as 250 plants since 1980 and have robbed American textile and apparel workers of hundreds of thousands of jobs during the past decade (*New York Times* 1985, p. 33). Indeed, the bill's drafters write that the total 1984 volume of textile and apparel imports represented "over 1 million job opportunities lost to the United States workers" (*New York Times*, p. 5).

Ellison McKissick, president of the American Textile Manufacturers Institute, argued that "our markets have been overwhelmed by

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

Richard B. McKenzie is Professor of Economics at Clemson University and Stephen D. Smith is Research Assistant at the Center for the Study of American Business at Washington University in St. Louis. The authors are indebted to Lee Benham, Kenneth Chilton, Arthur Denzau, Robert Parks, Chris Varvares, and Keith Womer for helpful comments on earlier drafts. The paper was written during Richard McKenzie's stay at the Center for the Study of American Business, where he was John M. Olin Visiting Professor of American Business.

[1]See Hudgins (1985) for a review of the history of textile and apparel import restrictions and of the features of the Textile and Apparel Trade and Enforcement Act of 1985.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

imports. . . . Since 1980, more than 300,000 fiber, textile and apparel workers have lost their jobs" (McKissick 1985, p. 23). And the backers of the 1985 textile and apparel bill argued: "If the rate of growth of imports of textiles and textile products into the United States that occurred since 1980 continues, plant closings will continue to accelerate, leaving the United States market with reduced domestic competition for imported products" (S.680, p. 8).

Concerned about plant closings in his home state caused partially by shirts that can be made in "downtown Shanghai, China for 18 cents an hour," South Carolina Senator Fritz Hollings called the textile and apparel bill "sensible protection." He and other supporters of protection maintained that all they want is to return to "fair trade, to make the table level again" (McKissick 1985, p. 23).

The purpose of this study is to assess the impact of changes in textile and apparel imports as well as domestic textile and apparel productivity on U.S. textile and apparel employment. While the findings are mixed, the research reported here should prove useful in future public debates over textile and apparel protection.

Textile and apparel employment combined would have fallen substantially (by possibly more than 200,000 jobs) during the 1973–84 period even if there had been no textile imports at all. Contrary to the contentions of protection proponents, textile imports have not in any systematic and predictable manner, or to any statistically significant extent, adversely affected U.S. textile employment between 1960 and 1985. However, apparel imports appear to have had a significant negative impact on employment in both industries.

Textile employment losses can, to a significant degree, be attributed to productivity improvements. (Determining the extent to which these productivity improvements may have been spurred by textile and apparel import competition is, however, beyond the scope of this paper.) On the other hand, productivity improvements have not been a statistically significant factor in employment losses in the apparel industry.

The findings of this study help explain why many textile firms have remained profitable in the face of employment cutbacks. Another, more predictable, conclusion of the analysis is that employment in both industries is significantly influenced by changes in real personal disposable income in the United States.

The paper begins with a brief description of developments in the textile and apparel industries between 1960 and 1984. A regression model is then used to analyze the effects of imports, personal disposable income, and productivity on textile and apparel employment.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The final section summarizes the results and conclusions drawn from the empirical findings.

## Industry Descriptions

In many respects, the textile and apparel industries are of similar size and have confronted similar employment, production, and import patterns during the 1960–84 period. During these two and a half decades, both industries faced mounting imports, productivity increases, and relative price declines.[2]

### *The Textile Industry*

Between 1960 and 1973 textile industry (SIC 22) employment rose from 924,000 to slightly more than 1 million, but by 1984 industry employment had fallen irregularly to 746,000, or by slightly more than a quarter of its 1973 peak employment level (see Figure 1).

Total industry shipments in constant-dollars (1984) nearly doubled between 1960 and 1973, rising from just over $29 billion to nearly $53 billion (see Figure 2).[3] After falling for two years, real shipments rose until 1978, peaking at nearly $65 billion. Textile shipments then began to fall again during the recessions of the early 1980s. By 1984, however, industry shipments had climbed back to just under $58 billion, in spite of a continuing drop in textile employment and an increase in textile imports.

During the 1960–84 period, worker productivity, defined as real shipments per worker, increased 147 percent, rising in real (1984) dollars from $31,407 per worker in 1960 to $77,526 per worker in 1984. Constant-dollar textile imports more than quadrupled during the period, yet market share expanded more modestly. Imports grew from $922 million, or 3.2 percent of total domestic textile shipments, in 1960 to nearly $3.8 billion, or nearly 6.5 percent of domestic textile shipments (see Figure 2).

How much have textile imports directly affected textile industry employment? A rough estimate of maximum potential employment can be obtained by assuming that textile industry employment is proportional to sales and then computing industry employment

[2]Employment data were taken from U.S. Bureau of Labor Statistics, *Employment, Hours, and Earnings: 1909–1984* and *Employment and Earnings* (July 1985); industry shipments and import data were taken from U.S. Department of Commerce, *U.S. Industrial Outlook* (various annual editions). The 1984 industry shipment and import figures are estimates.

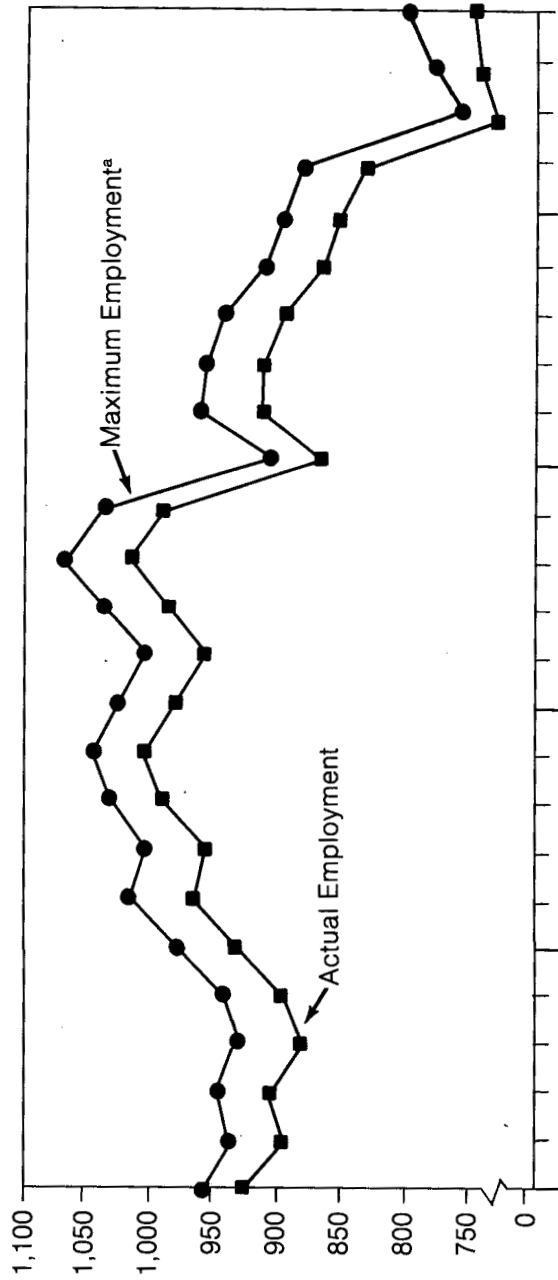
[3]Constant-dollar textile and apparel shipment and import figures were obtained by deflating the current-dollar figures by the textile and apparel price index (1984 = 100).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*FIGURE 1*

TEXTILE INDUSTRY EMPLOYMENT, ACTUAL AND MAXIMUM, 1960–84

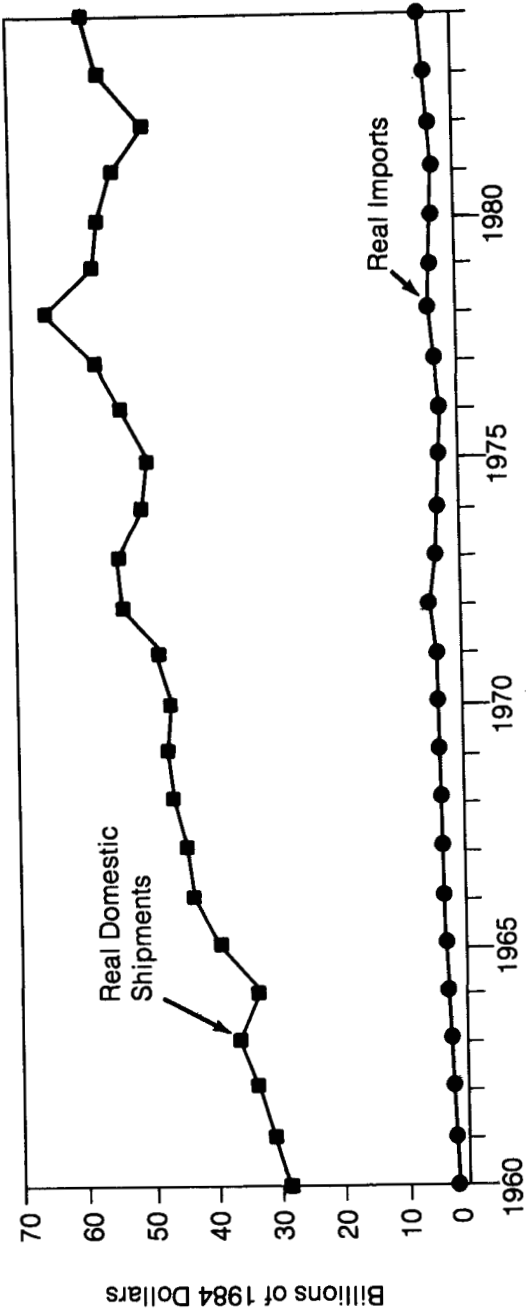



[a]Assumes imports are replaced by domestic production.

SOURCES: U.S. Bureau of Labor Statistics, *Employment, Hours, and Earnings: 1909–1984; Employment and Earnings* (July 1985).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*FIGURE 2*
TEXTILE INDUSTRY SHIPMENTS AND IMPORTS, 1960–84
(1984 Dollars)




SOURCE: U.S. Department of Commerce, *U.S. Industrial Outlook* (various annual editions).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

assuming that the import share would have been supplied by domestic firms.[4]

In the absence of imports, textile industry employment naturally would have been higher in every year between 1960 and 1984. (Compare the actual and maximum employment patterns in Figure 1.) Between 1973 and 1984, however, textile employment would still have fallen by virtually the same amount (267,000) as it actually did fall (264,000) even if imports could have been replaced by domestic production.[5] This is the case because the elimination of textile imports would have led to approximately the same increase in employment, about 50,000, in both 1973 and 1984.[6]

Practically the same conclusion can be reached about the impact of eliminating textile imports on textile employment between 1980 and 1984. During that period textile employment fell by 102,000, whereas industry employment would have fallen by 98,000 in the absence of imports. As will be discussed later, the substantial rise in real apparel imports, incorporating foreign textiles, did contribute, however, to declining employment in domestic textiles during both the 1973–84 and 1980–84 periods.

*The Apparel Industry*

In 1960 there were approximately 1.2 million apparel workers in the country. Employment in the industry (SIC 23) peaked in 1973 at over 1.4 million workers, fell to 1.2 million by 1975, and then rose to 1.3 million in 1979. In 1980 apparel employment began a steady decline to just under 1.2 million workers by 1984 (see Figure 3).

However, industry shipments in constant dollars moved irregularly upward from over $29 billion in 1960 to more than $57 billion in 1984 (see Figure 4). But unlike the textile industry, the value of industry shipments in apparel was greater in 1984 than in any previous year between 1960 and 1984.

Worker productivity doubled in real terms from $23,649 per worker in 1960 to $47,858 per worker in 1984. However, this increase was

[4]These estimates of the direct impact of textile imports are necessarily tentative because they do not account for other variables that can directly influence textile employment or the indirect effects of imports on employment through pressures on prices and productivity. The econometric work that follows partially remedies these problems.

[5]It is unlikely that domestic production could ever fully supplant imports. The higher prices of the domestically produced goods would cause total sales to drop below the level achieved with imports.

[6]In 1973 the elimination of imports, which represented 5 percent of domestic shipments, would have led to an increase in textile employment of about 50,000. In 1984 the elimination of imports, which then represented 6.5 percent of domestic shipments, would have led to an increase in employment of approximately 49,000.

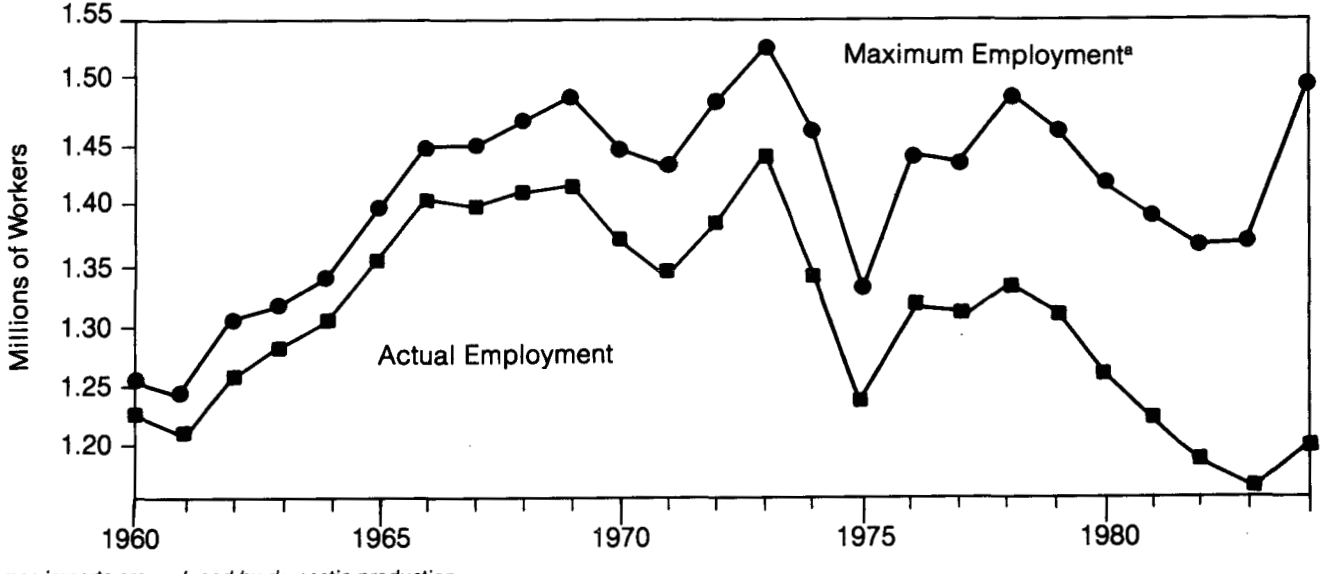


LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED



*FIGURE 3*
APPAREL INDUSTRY EMPLOYMENT, ACTUAL AND MAXIMUM, 1960–84




[a]Assumes imports are replaced by domestic production.

SOURCES: U.S. Bureau of Labor Statistics, *Employment, Hours, and Earnings: 1909–1984; Employment and Earnings* (July 1985).

slightly less than 70 percent of the productivity improvement in textiles. Constant-dollar apparel imports surged dramatically during the 1960–84 period, rising more than twentyfold from under $600 million, or 2 percent of domestic apparel sales, in 1960 to over $14 billion, or nearly 25 percent of domestic apparel sales, in 1984 (see Figure 4).[7]

Did the rise in apparel imports adversely affect apparel employment? If we again assume that apparel employment is proportional to sales and that apparel imports could be replaced by domestic production during the period 1960–84, apparel employment would have been 2 percent higher in 1960 and 25 percent higher in 1984.[8] The computed maximum apparel employment would have dropped 35,000 between 1973 and 1984, whereas actual employment fell by 241,000. In other words, between 1973 and 1984 apparel imports may have resulted in the loss of as many as 200,000 apparel jobs. (Compare the actual and maximum employment patterns in Figure 3.)

Much of the employment impact of apparel imports was indeed felt during the 1980–84 period, when U.S. apparel employment could have grown by 77,000 jobs but actually fell by 67,000. Possibly as much as three-quarters of the apparel job losses during the 1973 to 1984 period were caused by increased apparel imports over the first four years of the 1980s.[9]

*Prices of Textile and Apparel Products*

The 1984 price index for textile and apparel products was 111 percent above its 1960 level. However, the prices of other goods and services during the 1960–84 period rose much more rapidly, especially during the late 1970s. As a consequence, textile and apparel product prices declined relative to the consumer price index by one-third during the period.[10]

## Statistical Models and Results

The actual impact of textile and apparel imports on domestic employment can be more accurately assessed through regression

[7]Some of the surge in apparel imports could have been in response to quota restrictions on the importation of textiles.

[8]As with the textile industry, imports would have further reduced apparel employment through their indirect effects on industry prices and productivity.

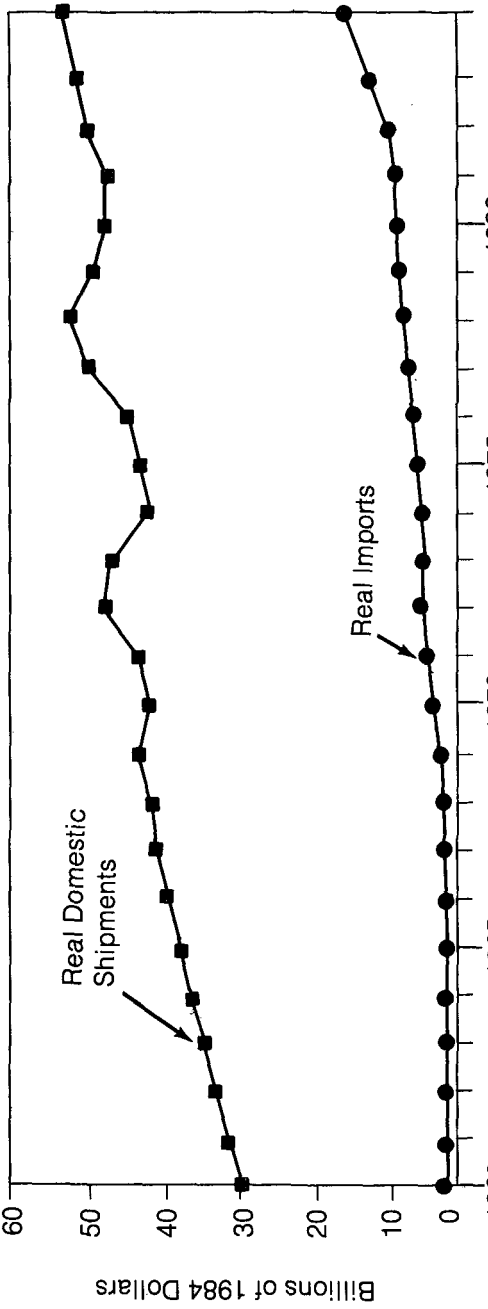
[9]Of course, lower-priced textile imports probably enabled domestic apparel firms to lower their costs and compete more effectively with apparel imports and kept apparel employment from falling as much as it otherwise would have.

[10]In 1984 the textile and apparel index (1967 = 100) stood at 211, whereas the consumer price index was 311.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*FIGURE 4*
APPAREL INDUSTRY SHIPMENTS AND IMPORTS, 1960–84
(1984 Dollars)




SOURCE: U.S. Department of Commerce, *U.S. Industrial Outlook* (various annual editions).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

analysis. Domestic textile and apparel employment can be influenced by many factors, including productivity change, real textile and apparel imports, and real disposable personal income. Textile and apparel employment may be expected to vary directly with real disposable personal income: the greater the real disposable personal income, representing buying power of consumers, the greater the demand for textile and apparel products and workers. The directional impact of the other three variables, however, is less certain. Productivity improvements can increase or decrease textile and apparel employment, depending on the elasticity of demand for textile and apparel products.

Textile and apparel imports can increase or decrease employment in these industries, depending on the relative magnitudes of the substitution and income effects of imports. Imports may substitute for U.S. firms' textile and apparel products because of either lower prices or higher quality.[11] In addition, imports—especially textile imports—can lower production costs of textile and apparel firms and prices of U.S. goods that use textile and apparel products, thus expanding sales and demand for U.S. or imported textile and apparel products.

The real income effects of textile and apparel imports can be expected to be positive; however, the substitution effects of imports—especially textile imports—are more ambiguous. On the one hand, apparel imports, ceteris paribus, can be expected to reduce domestic apparel and textile employment, while textile imports can make domestic apparel goods and textile goods that incorporate imported unfinished textiles more competitive. As a consequence, textile imports may have a positive effect on both domestic textile and apparel employment.[12]

To test the employment effects of imports, separate log-linear, least-squares regression models for the textile and apparel industries were developed, with total employment in the respective industries as the dependent variable. The time-series models covered annual

[11]Textile and apparel imports may also contribute to an expansion of real U.S. income via expanded purchasing power of consumer dollars that, in turn, may marginally expand the demand for U.S. textile and apparel products.

[12]Finally, an increase in textile and apparel imports may be caused by a rise in income and may accompany a rise in imports of a wide range of foreign goods and services (including textile and nontextile machinery and materials). Lower-priced and higher-quality imports of many goods and services can make American industries more competitive and can support the continued expansion of real income and employment in the U.S. economy. Of course, as noted in the text, textile and apparel imports can spur domestic firms to become more productive, which can indirectly lead to job losses in the textile and apparel industries.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

data for the period 1960–84. The four independent variables used in the regression equations are: (1) industry labor productivity as measured by the current-dollar value of total annual industry (textile or apparel) shipments adjusted by the textile and apparel price index divided by total industry employment;[13] (2) real disposable personal income as measured by current-dollar annual disposable income adjusted by the consumer price index; (3) real textile imports as measured by the current-dollar value of annual textile imports adjusted by the textile and apparel price index; and (4) real apparel imports as measured by the current-dollar value of annual apparel imports adjusted by the textile and apparel price index.

*The Effect of Imports*

The results of the least-squares analysis are presented in Table 1. As can be seen, the four variables explained nearly three-quarters of the variance in employment for both the textile and apparel equations. As indicated by their high F values, both equations offer robust explanations of textile and apparel employment over the 1960–84 period.

*Textile Imports.* During the 1960–84 period, real textile imports had a relatively small but statistically significant positive impact on both domestic textile and apparel employment levels, suggesting that textile imports may well have enabled some domestic producers to be more competitive through the use of cheaper unfinished textile imports. (The significance level of textile imports is lower for the apparel equation than the textile equation.) The two equations indicate that a 1 percent increase in real textile imports will lead to approximately a .16 percent increase in both the domestic textile and apparel employment.[14]

*Apparel Imports.* The coefficient for real apparel imports is negative and highly significant in both employment equations. The size of the negative apparel import coefficient, which is roughly the same in both equations, is also substantially higher than the positive coefficient for textile imports in both equations.

The statistical tests indicate that a 1 percent increase in real apparel imports will lead to approximately a .25 percent reduction in both

[13]The results of the empirical analysis should be evaluated with one caveat in mind. Since industry employment is used as the dependent variable and as the denominator in the measure of productivity, a measurement error for employment can introduce a negative bias in the coefficient for worker productivity.

[14]Conversely, the regression equation, using 1984 figures, suggests that reducing textile imports to zero theoretically could *reduce* textile employment by almost 125,000 jobs and apparel employment by almost 200,000 jobs.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 1*

LOG-LINEAR REGRESSION EQUATIONS FOR THE TEXTILE AND APPAREL INDUSTRIES

| Independent Variables | Dependent Variables: Total Employment in the Textile Industry | Dependent Variables: Total Employment in the Apparel Industry |
|---|---|---|
| Intercept | 0.731<br>(0.666)[a]<br>[.5131][b] | 2.222<br>(2.766)<br>[.0119] |
| Productivity | −0.455<br>(−2.123)<br>[.0464] | −0.076<br>(−0.400)<br>[.6937] |
| Real Disposable Personal Income | 1.143<br>(5.381)<br>[.0001] | 0.789<br>(5.659)<br>[.0001] |
| Real Textile Imports | 0.166<br>(2.243)<br>[.0364] | 0.159<br>(3.198)<br>[.0045] |
| Real Apparel Imports | −0.257<br>(−3.659)<br>[.0016] | −0.245<br>(−4.715)<br>[.0001] |
| Adjusted $R^2$ | .756 | .742 |
| F Value | 19.541<br>[.0001] | 18.266<br>[.0001] |

[a]Figures in parentheses indicate the t statistic.
[b]Figures in brackets indicate the significance level.

textile and apparel employment. However, it should be noted that if textile and apparel imports both expand by 1 percent, the increase in textile imports will partially offset the negative impact of the increase in apparel imports.[15]

*The Effect of Income*

As expected, during the 1960–84 period, real disposable personal income had a highly significant and strong positive effect on both textile and apparel employment. A 1 percent increase in real annual

[15]Put in more concrete terms, our statistical analysis, using 1984 data, indicates that reducing apparel imports to zero would increase textile employment by 191,000 and apparel employment by 293,000 jobs at most.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

disposable personal income can be expected to lead to a 1.14 percent increase in textile employment and a .79 percent increase in apparel employment. These findings reveal that, in terms of 1984 employment levels, a 1 percent increase in real disposable personal income can be expected to add approximately 8,500 textile jobs and 9,500 apparel jobs.[16]

*The Effect of Productivity*

Changes in worker productivity had different effects in the textile and apparel industries, perhaps reflecting a difference in the ability of the two industries to adjust to competitive pressures. Productivity changes did not have a statistically significant impact on apparel employment, whereas they did have a statistically significant and relatively strong negative influence on employment in the textile industry. A 1 percent increase in productivity in textiles can be expected to lead to a .46 percent reduction in textile employment.

*The 1973–84 Period.* As noted, textile employment peaked in 1973. Between 1973 and 1984, textile employment fell by 264,000 workers at the same time that worker productivity rose by 49 percent. According to the regression results of this study, that productivity increase had the potential of reducing textile employment by approximately 225,000 jobs (about 85 percent of the actual decrease in textile employment of 264,000).

*The 1980–84 Period.* The textile and apparel industries have been especially concerned with recent employment trends in their industries. The findings of this study suggest that the 21 percent increase in worker productivity in the textile industry during this period led to a decrease in textile employment of about 81,000 (or approximately 80 percent of the actual employment loss).

*Production and Peak Employment*

Some of the productivity increase experienced by the textile and apparel industries has resulted from the closing of inefficient plants, as well as from technological advances and competitive pressures brought on by imports. For this reason, the production capacities of these industries, for any given size of the labor force, must remain

[16]Future employment in the textile and apparel industries depends on the general health of the U.S. economy. If these estimates of the impact of income changes hold for the remainder of the 1980s, and if real personal disposable income rises by 3 percent a year throughout the remainder of the 1980s, textile and apparel employment combined could increase over 18 percent, or nearly 350,000 jobs, during the 1984–90 period. These combined employment gains in the textile and apparel industries can be expected to be tempered by productivity and import increases.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

rough estimates. Nevertheless, it is useful to note that, given the 1984 productivity of textile workers, if textile employment could be returned to the 1973 industry peak, textile output in 1984 would have been 35 percent greater, nearly $17 billion higher than the total of constant-dollar domestic production plus imports. Similarly, if apparel employment returned to the 1973 peak, the industry's output at current productivity levels would be 20 percent greater than it was, or only $3 billion less than the combined total of constant-dollar industry shipments and imports.

The point is that a return to the 1973 employment peaks in the textile and apparel industries is unlikely, even with drastic cutbacks in imports (as contemplated by the backers of the Textile and Apparel Trade Enforcement Act of 1985). Given the increased productivity in the two industries, the supply of products at the peak employment levels would approximately equal (in the case of apparel) or greatly exceed (in the textile situation) consumer demand.

*The Industry Impact of Imports*

The impact of textile and apparel imports may indirectly affect domestic textile and apparel employment through competitive pressure on prices and productivity. An expansion of imports can place downward pressures on textile and apparel prices and upward pressures on productivity. Those firms that lead the way in expanding productivity in the face of price declines can maintain—and even expand—employment and profitability. However, those firms that face declining relative prices without compensating increases in worker productivity can be forced to cut employment or close altogether.

Regression equations not reported here indicate that the relative price decline of textile and apparel products has had an impact on industry employment.[17] However, the exact amount of the price effect caused by import competition is unclear because the price, productivity, and employment effects are highly entangled. Textile and apparel prices can also be affected by domestic as well as foreign supplies that can be affected by productivity changes. In addition, the productivity changes can be a response to import pressures on prices.

[17]Indeed, when the ratio of textile and apparel prices to the consumer price index is introduced into each of the two equations as a fourth independent variable, the price ratio is positive, as might be expected, and highly significant. Imports are then statistically insignificant.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

## Summary and Conclusions

Several important conclusions can be drawn from this study. First, the findings cast doubt on the charge that textile imports are directly to blame for job losses in the textile and apparel industries. Job losses in the textile industry are not directly attributable to increases in textile imports during the early 1980s. If textile imports during the period had any negative effect at all on textile and apparel employment, it was likely an indirect effect through competitive pressures forcing U.S. textile firms to innovate and improve productivity.

The findings do, however, support apparel industry claims that employment has declined as a result of imports. In fact, both industries have been affected by apparel imports. The study also reaffirms the commonly acknowledged proposition that textile and apparel employment is strongly influenced by fluctuations in real disposable income.

Second, this study indicates that during the past 25 years, job losses in the textile industry have been due to a substantial degree to ongoing productivity increases. The same cannot be said about the apparel industry. One explanation is that the opportunities for labor-saving machinery to fend off import penetration appear to be much greater in textiles than in apparel, where labor-intensive "cutting and sewing" operations remain a significant part of the production process.

Third, the study suggests that much of the financial distress confronted by many domestic textile and apparel firms may be the consequence of expanded supplies of textile and apparel products caused in part by domestic productivity increases. Many firms that have closed or reduced operations have done so because they have been unable or unwilling to keep pace with industry productivity improvements. The analysis presented here helps explain why some major textile firms have been able to maintain rates of return on equity in excess of 10 percent (sometimes even significantly above 20 percent) at a time when textile and apparel plants are closing and employment is being cut.[18]

In summary, proponents of trade restrictions maintain that import protection is justified because of the presumed connection between observed textile and apparel imports and domestic employment. The

[18]In 1984, of the 21 major Southeastern textile firms evaluated, 12 had rates of return on equity in excess of 10 percent, of which 5 had rates of return on equity in excess of 15 percent. One firm had a rate of return on equity of more than 46 percent. Only two major textile firms on the list had losses in 1984 (*Jenks Southeastern Business Letter* 1985).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

presumed connection has been shown to be applicable in the case of apparel imports but not for textile imports. But in neither case is the magnitude of the employment loss anything approaching the one million lost jobs advertised by protectionists.

Further, as Will Rogers once said, "If a business thrives under a protective tariff, that don't mean that it has been a good thing. It may have thrived because it made the people of America pay more for the object than they should have, so a few got rich at the cost of the many." If apparel and textile imports are further restricted, the prices of such goods to consumers will likely rise, resulting in a hidden transfer of income from consumers to textile and apparel producers. Such curbs also will likely discriminate against lower-income groups because quantity controls typically cause a disproportionate reduction in lower-priced goods usually purchased by those groups.

Nor are the effects of quotas necessarily all positive for the textile and apparel industries themselves. Increased prices for domestic and imported textiles will increase the cost of textile and apparel production. The main long-run effect of curbs on textile and apparel imports seems to be some retardation of productivity improvements caused by reduced competitive pressures. Ultimately, slower productivity growth means that many U.S. textile and apparel firms will be less able to compete in the global marketplace.

## References

"Guilford Mills, Carriage Industries Rated Tops in Textile Survey." *Jenks Southeastern Business Letter.* 24 June 1985: 4–5.

Hollings, Fritz. "We're in a Trade War." *Reports.* Washington, D.C.: U.S. Senate Offices, July 1985.

Hudgins, Edwin. "Why Limiting Textile Imports Would Hurt Americans." *Backgrounder.* Washington, D.C.: Heritage Foundation, 30 September 1985.

McKissick, Jr., Ellison S. "Sweatshirts and Sweatshops." *Wall Street Journal* (4 September 1985): 23.

Megna, Laura, and Emrich, Thomas. "An Analysis of the Impact of the Textile and Apparel Trade Enforcement Act of 1985." Washington, D.C.: International Business and Economic Research Corporation, June 1985.

"Textile Industry Imports, Too." *New York Times* (10 October 1985): 5, 33.

U.S. Congress. Senate. "Textile and Apparel Trade and Enforcement Act of 1985." S.680, 26 February 1985.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# National Emergency and the Erosion of Private Property Rights

*Robert Higgs and Charlotte Twight*

> Only an emergency can justify repression.
> —Justice Brandeis

The scope of private property rights in the United States has been greatly reduced during the 20th century. Much of the reduction occurred episodically, as governmental officials took control of economic affairs during national emergencies—mainly wars, depressions, and actual or threatened strikes in critical industries. Derogations from private rights that occurred during national emergencies often remained after the crises had passed. A "ratchet" took hold. People adjusted first their actions, then their thinking, to accommodate themselves to emergency governmental controls. Later, lacking the previous degree of public support, private property rights failed to regain their pre-crisis scope.

Emergency restrictions of private property rights are by no means of concern only to historians of the growth of governmental power. Today, emergency restrictions limit many private rights, and many more sweeping restrictions could be lawfully imposed at the President's discretion. The possibility is real. Like several presidents before him, Ronald Reagan has dipped repeatedly into the government's reservoir of emergency economic powers. The potential exists for the greatly expanded use—and abuse—of such powers.

## Presuppositions

Rulers prefer more power to less, but in a liberal democracy the rulers are constrained by institutions that sustain private rights.[1]

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

Robert Higgs is the William E. Simon Professor of Political Economy at Lafayette College. Charlotte Twight is Assistant Professor of Economics at Boise State University and a member of the Washington State Bar.

[1]The ideas in this section are elaborated in Twight (1983, chap. 2), Higgs (1985), and Higgs (1987, chaps. 1–4).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Specifically, private property rights place the power of resource allocation in the hands of private citizens, thereby limiting the capacity of governmental officials to shape the economy. Governmental officials have interests of their own, which are not necessarily representative of or even in harmony with the interests of people outside government. Therefore, the rulers and the ruled normally struggle in various ways to determine who will control the use of resources. The greater the scope of private property rights, the more limited is the capacity of the rulers to achieve the ends they prefer at the expense of those preferred by the citizenry.

But citizens rely on the government for certain essential services, especially for the maintenance of social order and the protection of life and property; this dependence episodically creates opportunities for governments to take over previously private rights. If national emergencies did not just happen from time to time, governmental officials would be tempted to create them. In a genuine emergency, citizens are exceptionally willing to surrender their rights to governmental officials who offer plausible promises that they will restore social order, national security, or economic prosperity by wielding extraordinary powers. War, as Randolph Bourne aptly expressed it, is "the health of the state." Business depressions also promote robust government, as do nationwide strikes in strategic industries, especially those involving essential means of transportation or communication. Under modern ideological conditions—that is, if people insist that governments "do something" to rectify perceived socioeconomic problems—national emergencies invariably witness a transfer of economic rights from private citizens to governmental officials.

For several reasons, rights taken over by governmental officials during an emergency are unlikely to revert fully to their previous holders when normal times return. First, during the emergency the government learns how to operate its command-and-control system more successfully; that is, how to get necessary information, how to resolve competing claims on resources, and how to placate and respond to the complaints of politically influential aggrieved parties. Second, the imposition of a less-than-comprehensive system of controls—for example, regulating the economy but not the exercise of religious or political rights—discredits the conservative all-or-nothing warnings that normally inhibit governmental takeovers of private economic rights by representing them as steps toward totalitarianism. Third, many people discover, not only in the government's bureaus but in the more regulated "private" sector, that the controlled economy offers its own characteristic avenues to personal success. Those who



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

travel happily along these avenues naturally come to regard the entire system in which they thrive as essentially desirable. Some who adapt only out of necessity are eventually won over. Of course, during the emergency the government does all it can to justify its exercise of new powers, trumpeting the necessity and virtues of its controls, belittling the attendant costs and inequities, and defaming those who criticize its policies. Such propaganda is especially likely to hit its targets during a crisis, when the dogs of patriotism howl at their loudest and citizens rally more closely around the flag. Opponents of the government's emergency measures can be stigmatized as "slackers" or "draft-dodgers" or worse; in extreme cases, they may be imprisoned, deported, or deprived of normal civil liberties.

A great emergency, therefore, produces a ratchet: At an early stage the government, responding to an urgent and widespread insistence that it "do something," takes over rights previously held by private citizens; when the crisis wanes, public attitudes—the dominant ideology, some would say—have been so altered by the experience of governmental controls and the pervasive adaptations of behavior and thinking to those controls that public support for the recovery of the private rights is insufficient to produce their full restoration. While the hard residues of crisis-spawned laws, administrative agencies, and constitutional pronouncements are important, ultimately the most significant consequence of the emergency experience is the ideological change it fosters. As William Graham Sumner (1934, p. 473) wrote, "it is not possible to experiment with a society and just drop the experiment whenever we choose. The experiment enters into the life of the society and never can be got out again."

## How Emergencies Eroded Private Rights Historically

At the turn of the 20th century, Americans enjoyed a wide scope of private property rights.[2] The prevailing ideology of both elites and masses greatly emphasized economic liberty. As James Bryce (1895, pp. 536–37) observed, the typical American regarded the "right to the enjoyment of what he has earned" as "primordial and sacred." Most Americans believed that all governmental authorities "ought to be strictly limited" and "the less of government the better."

Federal officials and judges typically acted in conformity with the belief that government ought to be confined to protective functions,

[2]Documentation of the historical events described in this and the following section may be found in Higgs (1987, chaps. 5–10).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

refraining from redistributionist policies and leaving citizens free to conduct their economic affairs as they pleased. There were exceptions, of course, but they only highlight how limited the government, especially the federal government, was as a rule. In *Allgeyer v. Louisiana* (165 U.S. 578 [1897] at 589), the Supreme Court articulated the prevailing sentiment, declaring that the Constitution protects

> not only the right of the citizen to be free from the mere physical restraint of his person . . . but . . . the right of the citizen to be free in the enjoyment of all his faculties; to be free to use them in all lawful ways; to live and work where he will; to earn his livelihood by any lawful calling; to pursue any livelihood or avocation, and for that purpose to enter into all contracts which may be proper, necessary, and essential to his carrying out [these purposes] to a successful conclusion.

When citizens were guaranteed such extensive freedom of contract, governments had little ability to fix prices and wages, restrict hours or other conditions of employment, compel collective bargaining, regulate the locations of enterprises, or set aside the terms of private agreements; nor did governments have much justification for increasing the burden of taxation. Although governments, especially at the state and local levels, increased their economic functions somewhat during the late 19th and early 20th centuries and, more importantly, the dominant ideology gradually became more sympathetic to more-than-minimal government, the economy remained essentially a market system as late as 1914.

In a violent break with American tradition, private property rights were suppressed on a wide scale during World War I. In 1916, in anticipation of a future state of war, Congress granted the President emergency powers to seize materials, plants, and transport systems. During the year and a half of official U.S. belligerency that followed, the federal government took over the railroad, ocean shipping, telephone, and telegraph industries; suspended the gold standard and controlled all international exchanges of goods and financial assets; commandeered hundreds of manufacturing plants; entered into massive economic enterprises on its own account in such varied departments as shipbuilding, wheat and sugar trading, and building construction; lent huge sums to businesses directly or indirectly; regulated the private issuance of securities; established official priorities for the use of transport facilities, food, fuel, and many other goods; fixed the prices of dozens of important commodities; intervened in hundreds of labor disputes; and conscripted nearly three million men into the army. In short, the government extensively attenuated or destroyed private rights, creating what some contemporaries called



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

"war socialism." The Supreme Court did nothing to restrain the government's suppression of private rights; evidently, constitutional war powers covered all cases.

Although most private rights were restored after the war, not all were. Tax rates remained higher, the tax structure was more "progressive," and the income tax became a much more important source of federal revenues relative to traditional consumption taxes. The Transportation Act of 1920, by which the government relinquished its emergency control of the railroads, came close to nationalizing them; government remained in the ocean-shipping business; and the War Finance Corporation episodically participated in the credit markets until 1925. Most significantly, the war left a legacy of ideological change. As Bernard Baruch (1960, p. 74) observed, the war experience convinced many prominent businessmen and others that "government direction of the economy need not be inefficient or undemocratic, and suggested that in time of danger it was imperative."

The next "time of danger" was occasioned not by war but by the Great Depression, a national catastrophe that Justice Brandeis called "an emergency more serious than war." After three years of widespread bankruptcies, foreclosures, and bank failures, and of rapidly falling income and massively rising unemployment, it seemed that the market economy would never recover. All classes of Americans increasingly clamored for relief.

Recalling how the government had wielded sweeping emergency powers to manage the economy during the war, many politically influential people believed that similar governmental controls could be used effectively to "fight" the depression. In 1933, the federal government launched a fleet of emergency measures: huge work-relief projects; a program to cartelize virtually all industries; abandonment of the gold standard and prohibition of private domestic monetary transactions in gold; price and production controls in agriculture; detailed regulation of securities markets; extensive federal intrusion into labor markets and union-management relations; federal production and sale of electrical power; and, before the New Deal had spent itself in 1938, a multitude of federal insurance and credit programs, Social Security pensions and welfare payments, the minimum wage, national unemployment insurance, and many other forms of governmental intervention in the market economy.

For a while the Supreme Court resisted. The National Industrial Recovery Act and the first Agricultural Adjustment Act, centerpieces of the early New Deal, were declared unconstitutional. But the Court was of two minds. In 1934, in *Nebbia v. New York*, it upheld an emergency-inspired state law fixing milk prices and imposing criminal



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

sanctions on those who transacted at lower prices. The Court ruled (291 U.S. 502 [1934] at 523) that "neither property rights nor contract rights are absolute." In 1935, in the *Norman* case, sustaining the government's crisis-induced abrogation of the gold standard, the Court put private contractual rights in a clearly inferior position (294 U.S. 240 [1935] at 309–10): "There is no constitutional ground for denying to the Congress the power expressly to prohibit and invalidate contracts although previously made, and valid when made, when they interfere with the carrying out of the [monetary] policy it is free to adopt."

After vacillating during 1934–36, sometimes sustaining and sometimes striking down the government's unprecedented derogations from private property rights, the Court caved in completely in 1937. Since then it has maintained that virtually any state or federal governmental interference with private property rights is constitutional. Only a law that is manifestly arbitrary and lacking any imaginable relation to a public purpose will be disallowed. In a bloodless revolution, the U.S. Supreme Court overturned constitutional protections of private property rights that had existed for 150 years.[3]

The justices were only registering the ideological transformation going on around them. The Great Depression deeply discredited longstanding beliefs in individualism, private property rights, free markets, and limited government. The New Deal gave desperate people money, jobs, and protection from market competition, for which they were grateful. More importantly, it taught many people, including a new generation of Americans who had no personal experience with anything else, to value collectivist policies and governmental promises of economic security.

Drawing on the collectivist programs and sentiments of World War I and the New Deal, the federal government during World War II built an unprecedented apparatus of economic control. Clinton L. Rossiter (1948, p. 279) concluded: "Of all the time-honored Anglo-Saxon liberties, the freedom of contract took the worst beating in the war." Governmental authorities drafted 10 million men for service in the armed forces; allocated strategic materials; established priorities for the use of transport, food, fuel, and other goods; commandeered plants and sometimes whole industries; fixed prices and rents; rationed many consumer goods; and built and operated entire new industries. In short, the free market virtually disappeared during 1942–45.

[3]On the constitutional revolution of the late 1930s, see Murphy (1972), Siegan (1980), and Karlin (1983).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

While private property rights were being suppressed, the Supreme Court only approved. As Justice Douglas wrote in *Bowles v. Willingham* (321 U.S. 503 [1944] at 521), a case involving rent controls, "where Congress has provided for judicial review after the regulations or orders have been made effective it has done all that due process under the war emergency requires." But Justice Roberts, dissenting in a related case, *Yakus v. United States* (321 U.S. 414 [1944] at 458), correctly described the judicial review as "a solemn farce." The government simply did as it pleased, showing no regard for private property rights. As Rossiter (1976, p. 91) commented, "the Court, too, likes to win wars." Even when the government herded some 110,000 persons of Japanese descent, two-thirds of them U.S. citizens, into concentration camps, the Court stood aside, declaring the peremptory imprisonment of these innocent persons "in the crisis of war" to be "not wholly beyond the limits of the Constitution" (*Hirabayashi v. United States*, 320 U.S. 81 [1943] at 101). Where, one wonders, *are* the constitutional limits on governmental invasion of private rights during emergencies? During World War II, not even the most elementary private rights could withstand the government's determination to exercise emergency powers.

Enduring legacies of the war include the Employment Act of 1946, which committed the federal government to a policy of ongoing macroeconomic management; the Taft-Hartley Act of 1947, which preserved many of the federal powers first stipulated in the War Labor Disputes Act of 1943; and the Selective Service Act of 1948, which extended into peacetime the military conscription by which the government had coercively obtained the bulk of its military manpower at below-market rates during the war. The emergency rationale held sway. As a congressman (quoted by Gillam 1968, p. 509) said during the debate on the peacetime draft bill, "What we propose to do today is, under ordinary conditions, contrary to our traditions. . . . We are not living in an ordinary time. . . . We are living in a world of fear, of chaos, of uncertainty. . . ."

Most significantly, as Calvin B. Hoover (1959, p. 212) observed, the war experience "conditioned [businessmen] to accept a degree of governmental intervention and control after the war which they had deeply resented prior to it." Even under the pro-business Eisenhower administration, no serious attempt was made to overthrow the economic controls inherited from past emergencies. Having so greatly adjusted their actions and their thinking to the emergency suppression of private property rights, most Americans seemed to fear a return to a free market regime more than they feared the denial of private rights, actual and potential, under the postwar politico-economic arrangements.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

## A Case in Point: The Labor Market

The inverse relation of national emergency and private property rights seems clear enough. Can a closer inspection of the historical record sustain the same conclusion? We think so. Elsewhere we have described and analyzed in much greater detail the economic policies adopted during the world wars and the Great Depression (see Higgs 1987, chaps. 7–9), and we have documented the emergency ratchet with regard to private rights in international trade, a glaringly clear case (see Twight 1985).

Here we shall extend our examination of the historical record by looking at the emergency ratchet as it pertains to private rights in labor markets. Before World War I, private transactors possessed extensive freedom of contract in labor markets. Today, such rights are greatly attenuated by, inter alia, many federal restrictions: minimum wage law, wage-and-hour laws, compulsory collective bargaining laws, retirement and pension laws, anti-discrimination laws, and others. How did the American people get from there to here?

So far as federal restrictions are concerned (state and local restrictions have a much longer and more complicated history), the first significant step was taken in 1916 with the passage of the Adamson Act. The railroad-operating brotherhoods had demanded a reduction of the standard working day from 10 to 8 hours without any reduction in daily pay. Caught between the unions' impending increase of railroad labor costs and the Interstate Commerce Commission's (ICC's) stubborn restraint of railroad rates and fares, the employers refused to accept the proposal; the unions scheduled a nationwide strike. The prospect of such a calamitous work stoppage terrified President Wilson, who decided to intervene. Failing to bring the two sides into agreement, Wilson induced Congress, which was controlled by his fellow Democrats, to enact legislation establishing the eight-hour day with daily pay unchanged, that is, mandating a 25 percent increase in the wage rates of operating employees of interstate railroad companies (39 *Stat.* 721, 3, 5 Sept. 1916). With this legislative action, as Jonathan R. T. Hughes (1977, pp. 138–39) wrote, the nation began a "long and irregular movement away from the ideal of the wage bargain as a sacred bond between the employer and the individual employee"—in particular, a movement away from federal support for private rights of contract in the labor market.

The emergency of the threatened railroad strike in the winter of 1916–17 merged immediately into the emergency of economic mobilization after the American declaration of war the following April. Unsatisfied with their recently mandated wage increase, the railroad-operating brotherhoods renewed their threat to strike toward the end



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of 1917. The railroad companies, still caught between the unions' wage pressure and the ICC's rate restraint (with threats of antitrust prosecution further complicating the situation and impeding cooperation among the companies) struggled ever more unsatisfactorily to meet the extraordinary demands for service occasioned by the government's massive, ill-coordinated procurement program. Traffic slowed and facilities grew increasingly snarled. Knowing no other way to break the impasse, which was almost entirely of the government's own creation, President Wilson took over the railroad industry under the powers granted to him by the Army Appropriations Act of 1916. "It was done," wrote Treasury Secretary William G. McAdoo (1931, p. 458), only "as an imperative war measure." McAdoo, who became the Director-General of the Railroad Administration, soon ordered substantial increases of railroad wage rates. The Railroad Administration's resulting losses were conveniently covered by drawing on the U.S. Treasury.

Beyond the railroad industry, the government faced the same threat that strikes might cripple its mobilization of resources for war purposes. During 1917, there were an unprecedented 4,450 work stoppages (U.S. Bureau of the Census 1975, pp. 178–79). To allay the threat, the Wilson administration adopted a policy of supporting union organization and compulsory collective bargaining, the eight-hour day, and union work rules. If necessary, employers could be indemnified by cost-plus contracts with governmental procurement agencies. In 1918, a National War Labor Board was created to intervene in labor-management bargaining and bring the collective agreements into forms acceptable to the government. In extreme cases, the President and his lieutenants threatened to draft recalcitrant unionists or to commandeer plants owned by recalcitrant employers. Legacies of the government's wartime labor policies included the Railroad Labor Board, an agency created by the Transportation Act of 1920, and the Railway Labor Act of 1926 and of 1934.

When the Great Depression of the early 1930s prompted the Roosevelt administration to devise an emergency industrial relief program, the labor provisions of the Wilsonian war policy reappeared—a sop to unionists to keep them from opposing the cartels being established for producers. Along with clearly specified collective bargaining rights, the labor provisions of the National Industrial Recovery Act (48 *Stat.* at 198–99) provided "that employers shall comply with the maximum hours of labor, minimum rates of pay, and other conditions of employment, approved or prescribed by the President." So much for freedom of contract in the labor market. The Supreme Court invalidated the NIRA early in 1935, but Senator



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Robert Wagner and his congressional colleagues quickly enacted the National Labor Relations Act to preserve governmental support for compulsory collective bargaining. Three years later the Fair Labor Standards Act established minimum-wage and wage-and-hour provisions similar to those lost when the NIRA was struck down.[4]

The next great emergency, World War II, brought much greater attenuation of private rights in the labor market. Ten million men were drafted into the armed forces. A War Manpower Commission made various attempts to impede labor mobility, reallocate workers, and control hiring practices; at one point it issued a "work-or-fight" order, threatening men in "nonessential" jobs with conscription. The National War Labor Board intervened extensively in collective bargaining negotiations, generally supporting union demands for representation and backing up its authority with dozens of seizures, some involving entire industries (coal mining and, no surprise, railroads). Irritated by strikes in strategic industries and most of all by John L. Lewis's plain defiance of the President, Congress enacted the War Labor Disputes Act of 1943. This law, inter alia, expanded the President's power to seize production facilities. (The Taft-Hartley Act of 1947 drew heavily on the provisions of the wartime labor law.)

Proposals for compulsory universal service received serious consideration during 1943–45 (FDR himself favored such a system)but Congress never enacted any of the proposed bills. Military conscription, on the other hand, became firmly entrenched. After the war, "temporary, emergency" extensions of the draft followed in succession. Not until the early 1970s did the government abandon its use of coercion to secure the manpower deemed necessary for national defense. Today's registration requirement makes clear, however, that the government stands ready to reinstate the military draft whenever political authorities deem it expedient. Young men no longer hold a constitutionally guaranteed right of ownership over their bodies, their labor power, or their lives.

It would not be difficult to relate the additional federal attenuation of private labor-market rights during 1964–74 (Civil Rights Act, Age Discrimination in Employment Act, Economic Stabilization Act of 1970, Occupational Safety and Health Act, Equal Employment Opportunity Act, Employee Retirement Income Security Act) to the

[4]In the interest of brevity, we only mention the New Deal's massive work-relief programs, emergency measures that established a permanent precedent for the federal government to function as an employer of last resort. Legacies include the Employment Act of 1946, the Comprehensive Employment and Training Act, and other such intrusions in the labor markets.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

quasi-emergency conditions that prevailed during the Great Society-Vietnam War era, a time of race riots, antiwar protests, massive public demonstrations, and political assassinations. But enough has already been shown to establish a clear historical link between national emergency and the attenuation of private property rights in the labor market. The landmark federal intrusions in this area occurred when governmental officials undertook to "do something" to deal with conditions widely regarded as national emergencies—great strikes, war mobilization, deep depression. And doing something almost always meant transferring rights from private citizens to governmental officials.

## How Emergency Powers Continue to Be Exercised

Until the late 1970s, unrevoked presidential declarations of national emergency, left over from long-past crises, continued to sustain extraordinary governmental authority. As late as 1976 the national emergencies declared by Roosevelt in 1933 and by Truman in 1950 had not been terminated. After the end of the Korean War, even though economic and political conditions in the United States were often as normal as possible in an age of Cold War and nuclear weapons, the government continued to use a fictitious emergency rationale to augment its statutory power—and thereby to diminish the scope of private property rights.

Few people worried about the emergency-spawned powers until the Nixon years. President Nixon himself added to the perpetual official crisis by declaring national emergencies in order to deal with the postal workers' strike in 1970 and the balance-of-payments "crisis" in 1971. These declarations remained unrevoked when their precipitating crises passed. Only when the Watergate scandal and Nixon's impoundment of congressionally authorized funds engendered widespread fear of an imperial presidency did Congress begin to investigate seriously the web of governmental authority spun on the pillars of unrevoked executive declarations of national emergency.

In 1973, the Senate created a Special Committee on the Termination of the National Emergency (subsequently redesignated the Special Committee on National Emergencies and Delegated Emergency Powers) to investigate the matter and to propose reforms. Ascertaining the continued existence of four presidential declarations of national emergency, the Special Committee (U.S. Senate 1973, p. iii) reported:

> These proclamations give force to 470 provisions of Federal law. . . . taken together, [they] confer enough authority to rule the country



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> without reference to normal constitutional processes. Under the powers delegated by these statutes, the President may: seize property; organize and control the means of production; seize commodities; assign military forces abroad; institute martial law; seize and control all transportation and communication; regulate the operation of private enterprise; restrict travel; and, in a plethora of particular ways, control the lives of all American citizens.

Yet, there was no statutory definition of what constitutes an emergency; as with the fabled emperor's clothes, the assertion could belie the reality. As its investigation proceeded, the committee repeatedly voiced concern about both the longevity of the claimed emergencies and the use of a single national emergency decree to effectuate widely scattered emergency powers wholly unrelated to the emergency at hand.

The ultimate outgrowth of the committee's work was the enactment of two significant laws: the National Emergencies Act (NEA) of 1976 and the International Emergency Economic Powers Act (IEEPA) of 1977.[5] The NEA represented an attempt to reassert congressional control over burgeoning national emergency authority. Terminating all existing national emergencies two years after its enactment, it permits future presidential declarations of national emergency only under strict procedures guaranteeing periodic executive and congressional review of each emergency decree. The act also requires that no statutory authority contingent on a state of emergency be triggered by a national emergency decree unless the President specifically states his intent to exercise authority under that particular statute. Moreover, the act declares that national emergencies may be terminated by concurrent resolution of Congress as well as by presidential proclamation or automatic lapse if not renewed annually by the President.

As originally enacted, the NEA *exempted* Section 5(b) of the Trading with the Enemy Act (TWEA) from its provisions, simultaneously mandating immediate congressional review of national emergency powers and programs under that authority. The exemption was significant because since 1917 the TWEA had given the President virtually unlimited power during war or (since 1933) other national emergencies to regulate or prohibit transactions in foreign exchange or international credit as well as purely domestic banking activities. As Representative Jonathan Bingham (U.S. House of Representatives

[5]National Emergencies Act, Pub. L. 94-412, 90 *Stat.* 1255 (Sept. 14, 1976), codified as 50 *U.S.C.*, Sections 1601–51 (1976); International Emergency Economic Powers Act, Pub. L. 95-223, 91 *Stat.* 1625 (Dec. 28, 1977), codified as 50 *U.S.C.*, Section 1701 et seq. (1976 ed., Supp. V, 1981).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

1977a, p. 5) later described it, "the TWEA was so broad that the President could do virtually what he wanted under it under any kind of an emergency and it did not have to involve war or hostility or anything else." The congressman believed these powers to be "enormously broad" and potentially "dictatorial" and worried that they could be employed "without any restraint by the Congress."

In the 1977 TWEA committee hearings, the clash between members of Congress seeking to control this executive power and officials of the executive branch intent on sustaining it became intense. Incredulous congressmen learned the full reach of executive encroachments on private property rights undertaken in the name of obsolete emergency decrees, as the exchange between Representative Bingham and State Department official Julius Katz illustrates (U.S. House of Representatives 1977b, p. 110):

> Mr. BINGHAM. Mr. Katz, what is the national emergency currently facing us which warrants the use of powers under the Trading with the Enemy Act? . . . What is the national emergency today?
>
> Mr. KATZ. It continues to be the emergency involving the threat of Communist aggression which was declared in 1950 at the time of the aggression in Korea.
>
> Mr. BINGHAM. *Are you serious?*

Moments later, referring to "this 'Alice in Wonderland' situation that we are in today with the Trading With the Enemy Act," Bingham (pp. 111–12) complained that

> the administration wants to continue exercising powers which in the mind of any fair-minded observer . . . are totally unrelated to the purposes of the act. We don't know who the enemy is that we are talking about. We don't have an emergency that is current. . . . we are playing fast and loose with the Constitution. We are relying on an emergency that does not exist to justify a taking, or actions taken by the administration.

Ultimately, even administration witnesses acknowledged the historical vulnerability of private rights to the political imperatives of crisis as reflected in the evolution of the TWEA. Treasury official C. Fred Bergsten, a key administration spokesman on the issue, said that he was "keenly aware" that "on several occasions section 5(b) has been hurriedly broadened during moments of national crisis, such as the banking emergency in 1933 and World War II breakout in 1941, during which cases very little attention, frankly, was given to procedural safeguards consonant with the constitutional balance of power" (U.S. Senate 1977, p. 2).

To narrow the TWEA's exemption from the NEA, Congress passed the IEEPA in 1977. Viewed broadly, the effect of its enactment is to



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

separate presidential power over transactions previously covered by Section 5(b) into two categories. The TWEA is amended to authorize presidential action exclusively during wartime, while the IEEPA becomes a repository for executive economic power wielded during other national emergencies. Significant also is the IEEPA's restriction that the emergency powers it authorizes apply to international transactions only. Withheld are the purely domestic emergency powers previously exercised under the all-inclusive TWEA.

Presidential emergency authority over international transactions, however, remains virtually unchanged. In language parallel to the wartime provisions of the TWEA, the IEEPA grants the President emergency authority to

> (A) investigate, regulate or prohibit—
> (i) any transactions in foreign exchange,
> (ii) transfers of credit or payments between, by, through, or to any banking institution, to the extent that such transfers or payments involve any interest of any foreign country or a national thereof,
> (iii) the importing or exporting of currency or securities; and
> (B) investigate, regulate, direct and compel, nullify, void, prevent or prohibit, any acquisition, holding, withholding, use, transfer, withdrawal, transportation, importation or exportation of, or dealing in, or exercising any right, power, or privilege with respect to, or transactions involving any property in which any foreign country or a national thereof has any interest;
> by any person, or with respect to any property, subject to the jurisdiction of the United States.

These copious powers are to be used "to deal with any unusual and extraordinary threat, which has its source in whole *or substantial part* outside the United States, to the national security, foreign policy, *or economy* of the United States, if the President declares a national emergency with respect to such threat."[6]

Although the President's emergency authority over international transactions retained its previous scope, the power to effect the requisite emergency was narrowed. In enacting the IEEPA, Congress repealed the exemption that had shielded Section 5(b) of the TWEA from the NEA. The IEEPA thus guaranteed that *future* declarations of national emergency within its scope would be subject to the NEA and that current "emergencies" could be terminated as specified in that statute. Nevertheless, *existing uses* of Section 5(b) powers continued to be insulated from the NEA by provisions enabling the President to extend current Section 5(b) programs annually upon

[6]50 *U.S.C.*, Section 1701(a) (1976 ed., Supp. V, 1981), emphasis added.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

executive determination that such extension "is in the national interest"—that is, without the necessity of a continuing national emergency. Existing Section 5(b) programs included the Foreign Assets Control Regulations (restricting American citizens' rights to enter into trade or financial transactions with citizens of North Korea, Vietnam, Cambodia, and, to a lesser extent, China), Cuban Assets Control Regulations, Transactions Control Regulations (prohibiting U.S. citizens from engaging in certain transactions involving strategic goods of foreign origin), and Foreign Funds Control Regulations (affecting assets of citizens of East Germany, Czechoslovakia, Latvia, Lithuania, and Estonia that were blocked during World War II).

How well have the NEA and the IEEPA served their intended purpose? In the decade since their enactment, the attempted curtailment of executive power has proved to be significantly less successful than the statutory language suggested it would be.

In one sense, the dramatic and frequent use of executive emergency power that persists under the two statutes speaks for itself. President Carter and President Reagan have declared national emergencies repeatedly to impose sweeping economic restrictions on the dealings of Americans with citizens of other nations, including Iran, Nicaragua, South Africa, and Libya. Most recently, invoking the NEA and the IEEPA, President Reagan declared a national emergency to warrant a prohibition of the economic activities of Americans in Libya and ordered all U.S. citizens residing in Libya to leave or face criminal sanctions. Also under claim of national emergency, President Reagan recently used the IEEPA to sustain an important trade restriction, the Export Administration Act, for more than a year during which Congress had allowed it to expire.

The Libyan episode is particularly revealing of the fundamental issues involved in recent emergency declarations. Although the President and the popular press invite us to view the restrictions as impinging on the Libyan government, in reality the emergency decrees restrict the rights of American citizens. In his Executive Order 12543 of January 7, 1986 (51 *Fed. Reg.* 875, Jan. 9, 1986), President Reagan explicitly forbade U.S. citizens to engage in

> (a) The import into the United States of any goods or services of Libyan origin. . . ;
>
> (b) The export to Libya of any goods, technology (including technical data or other information) or services from the United States. . . ;
>
> (c) Any transaction by a United States person relating to transportation to or from Libya; . . . or the sale in the United States by any person holding authority under the Federal Aviation Act of any transportation by air which includes any stop in Libya;



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> (d) The purchase by any United States person of goods for export from Libya to any country;
>
> (e) The performance by any United States person of any contract in support of an industrial or other commercial or governmental project in Libya;
>
> (f) The grant or extension of credits or loans by any United States person to the Government of Libya, its instrumentalities and controlled entities;
>
> (g) Any transaction by a United States person relating to travel by any United States citizen or permanent resident alien to Libya, or to activities by any such person within Libya, after the date of this Order, other than transactions necessary to effect such person's departure from Libya. . . ; and
>
> (h) Any transaction by any United States person which evades or avoids, or has the purpose of evading or avoiding, any of the prohibitions set forth in this Order.

Whatever one's views of the events that prompted the President's actions, one cannot escape the conclusion that the economic activities of *Americans*, not Libyans, are controlled directly by the emergency restrictions. It became evident that the valuable investments of Americans in plant and equipment stood a good chance of being relinquished to the Libyan government, without compensation, by the terms of the President's executive order (see Kempe 1986a).

Caught in its own snare, the Reagan administration changed the program, making exceptions to and deferrals of its original orders in hopes of avoiding the enormous financial gains to Qadhafi's regime that rigid enforcement of the edicts would create (see Greenberger 1986 and Kempe 1986b). While observing the government's confused thrashing, one might well have reflected on who was supposed to suffer as a result of the emergency economic sanctions and on whether or not such suffering could reasonably be expected to result from such restrictions of American rights. These events prompt one to wonder about what events might provoke a charismatic President to impose sanctions that selectively restrict the property rights of an unpopular subgroup of Americans—such as the Japanese-Americans in 1942—under the pretext of emergency.

## The Constitutionality of Emergency Powers

Are emergency powers constitutional? The short answer is yes. One sense of "constitutional" is whatever the government makes a practice of doing. Since the federal government took emergency action to head off the threatened railroad strike in 1916, it has repeat-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

edly exercised emergency powers. To conduct a practice for 70 years is to establish that it is not simply an aberration.

Another meaning of "constitutional" is whatever accords with the U.S. Constitution. Ours is a written document that completely enumerates the powers that may be exercised by the federal government. The Tenth Amendment guarantees that powers not explicitly assigned to the central government nor prohibited to the state governments belong to the states and the people. The Constitution makes no provision for the exercise of emergency powers. Evidently, it was meant to govern the actions of federal authorities in foul weather as well as in fair. How, then, has it been possible for federal officials to exercise—eventually as a matter of routine—extraordinary powers for which the Constitution provides no warrant?

The answer lies in a third meaning of "constitutional"; that is, whatever the Supreme Court allows. The Court has ruled on several occasions on the permissibility of emergency powers; these decisions constitute a melancholy chapter of constitutional history, a record of evasion and capitulation of the judicial function against which many justices on the minority side have objected.

The first modern ruling, still a leading precedent, was handed down in *Wilson v. New* (243 U.S. 332 [1917]), a case arising from a challenge to the government's emergency appeasement of the railroad unions in 1916. By a 5-4 margin, the justices upheld the government's actions. Speaking for the Court, Chief Justice White (pp. 350–51) argued that the Adamson Act was a legitimate exercise of the legislative will "to the end that no individual dispute or difference might bring ruin to the vast interests concerned in the movement of interstate commerce. . . ." The dispute, he observed, "if not remedied, would leave the public helpless, the whole people ruined and all the homes of the land submitted to a danger of the most serious character."

The Chief Justice (pp. 348, 352) repeatedly emphasized the gravity of the circumstances prompting the Adamson Act: "the impediment and destruction of interstate commerce which was threatened" and "the infinite injury to the public interest which was imminent." Oddly, he took pains to *deny* that the emergency *per se* gave rise to the government's authority under the act: "although an emergency may not call into life a [constitutional] power which has never lived," he reasoned, "nevertheless emergency may afford a reason for the exertion of a living power already enjoyed." In view of the unprecedented character of the government's actions, White's distinction was a most delicate one; not everyone could see the sense of it.

The dissenting justices declared that emergency conditions, however threatening, could not excuse the denial of constitutional rights.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In their view, the Constitution had been adopted to protect private property *especially* under such conditions. "The suggestion," wrote Justice Pitney (pp. 376–77), that the Adamson Act "was passed to prevent a threatened strike . . . amounts to no more than saying that it was enacted to take care of an emergency. But an emergency can neither create a power nor excuse a defiance of the limitations upon the powers of the Government." Despite their strong arguments, the dissenters could do no more than register their dissatisfaction. The majority's curious doctrine, that emergencies may call forth extraordinary powers from some previously untapped yet still legitimate reservoir, carried the day and established the precedent.

A broad construction of the war powers excused the government's suppression of private rights during World War I, but legal challenges persisted into the postwar era. In the spring of 1921, the Supreme Court ruled that a rent-control ordinance that Congress had imposed in late 1919 in the District of Columbia—"its provisions were made necessary by emergencies growing out of the war"—"was not, in the prevailing circumstances, an unconstitutional restriction of the owner's dominion and right of contract or a taking of his property for a use not public." The ruling, which emphasized the temporariness of the disputed rent controls, seemed to the dissenters to open the door to all kinds of governmental restrictions of the private right of contract. "As a power in government," wrote Justice McKenna, "if it exist at all, it is perennial and universal. . . . necessarily, if one contract can be disregarded in the public interest every contract can be. . . . other exigencies may come to the Government making necessary other appeals."[7] Indeed they would.

Early in 1934 the Court issued what is perhaps the classic decision on the emergency powers of government and the protection to be afforded or withheld from private property rights during economic crises (*Home Building and Loan Association v. Blaisdell*, 290 U.S. 398 [1934]). The case involved a legislative moratorium—one of 25 such state laws enacted during 1932–33—on mortgage foreclosures

[7]Block v. Hirsch, 256 U.S. 135 (1921) at 136, 168–69. Also Marcus Brown Holding Company, Inc. v. Feldman et al., 256 U.S. 170 (1921). Not until 1923 did the Supreme Court decide that the emergency had passed and refuse to validate a 1921 extension of the 1919 rent-control act in the District of Columbia (see Chastleton Corp. v. Sinclair, 264 U.S. 543 [1924]). Commenting on this decision, Edward S. Corwin (1947, p. 83) wrote: "The power of the Court to take notice of an emergency is, it appears, a two-edged sword, but its anti-government edge descends only after the emergency is over." Fifty years after the court ruled on *Block* and *Marcus Brown*, these decisions were employed to establish the constitutionality of a Rent Control Enabling Act when the government of Cambridge, Massachusetts, declared a "housing emergency" and imposed local rent controls. See Navarro (1984, pp. 14, 302).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

(see Alston 1984). A statute approved by the government of Minnesota on April 18, 1933, under ominous pressures by indebted farmers, declared an economic emergency and extended temporarily the period during which creditors were prevented from foreclosing on and selling mortgaged real estate. At the discretion of local courts the owners of foreclosed property could be allowed up to two years to redeem their property. During the extended redemption period the mortgagor was required to pay the mortgagee what amounted to rent, but the mortgagee was deprived of the right to take control over the property. The Home Building and Loan Association of Minneapolis challenged the law as an unconstitutional impairment of the obligation of contract. Its position was upheld by a county court but overturned by the state supreme court; the case then came before the U.S. Supreme Court on appeal.

The main issue was whether emergency conditions justified the exercise of otherwise unconstitutional powers by a state government. The most compelling precedents for an affirmative answer were *Wilson v. New* and the 1921 rent-control cases. The force of the rent-control decisions was uncertain, however, as Justice Holmes had previously declared that "they went to the very verge of the law" (*Pennsylvania Coal Company v. Mahon,* 260 U.S. 393 [1921]). Counsel for the state of Minnesota conceded (290 U.S. at 409) that "in normal times and under normal conditions" the state's moratorium law would be unconstitutional. "But these," he urged, "are not normal times nor normal conditions. A great economic emergency has arisen in which the State has been compelled to invoke the police power. . . ." The justices would have to decide again whether or not an emergency altered the protections of the Constitution and, if so, what qualified as an emergency. Their decision could have wide repercussions, because virtually all of the important federal statutes enacted during Franklin D. Roosevelt's first 100 days had appealed to emergency conditions, evidently with an eye toward the Supreme Court (see Clark 1934 and Belknap 1983).

Although the Court upheld the Minnesota moratorium law and by implication all the others, it failed to clarify whether or not an emergency justified setting aside the constitutional protection of private property rights. Perhaps the most important aspect of the decision is simply that state infringement of contractual obligations was validated. Paradoxically, however, the Court *denied* that it had sustained the Minnesota statute because of emergency (290 U.S. at 425, 426, 428, 437, 440–41, 444). "Emergency does not create power," said Chief Justice Hughes, speaking for the five-man majority. "Emergency does not increase granted power or remove or diminish the



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

restrictions imposed upon power granted or reserved." What could be plainer? Yet, wrote Hughes, referring to the slippery doctrine enunciated in *Wilson v. New*, "while emergency does not create power, emergency may furnish the occasion for the exercise of power." The prohibition of governmental intervention guaranteed by the Contracts Clause, he added, "is not an absolute one and is not to be read with literal exactness like a mathematical formula. . . . The economic interests of the State may justify the exercise of its continuing and dominant protective power notwithstanding interference with contracts." Any doubts about the capaciousness of the police powers of the state government and the implied limitation on the protection of private contractual rights, he believed, should have been removed by the rent-control decisions, where "the relief afforded was temporary and conditional . . . sustained because of the emergency." The Chief Justice insisted that "the reservation of the reasonable exercise of the protective power of the State is read into all contracts. . . ." He concluded: "An emergency existed in Minnesota which furnished a proper occasion for the exercise of the reserved power of the State to protect the vital interests of the community."

The four dissenting justices disagreed emphatically (pp. 448, 465, 471). "He simply closes his eyes to the necessary implications of the decision," said their spokesman, Justice Sutherland, "who fails to see in it the potentiality of future gradual but ever-advancing encroachments upon the sanctity of private and public contracts." The correct doctrine, they believed, was to be found in the classic decision of *Ex parte Milligan* (1866), which declared constitutional guarantees and restrictions to be absolutely invariant with respect to emergencies or any other social conditions. Quoting from a considerable collection of historical works, Sutherland established that the framers had intended the Contracts Clause to apply *"primarily and especially"* during economic crises; a crisis similar to the current depression had provoked them to add it to the Constitution in the first place. "The present exigency is nothing new," Sutherland pointed out. The rent-control decisions were dismissed as too weakly justified at the time and as dealing with a matter too dissimilar to the present one to afford a binding precedent.

The dissenters on the Court rejected Hughes's sophistical distinction between genuine emergency powers and reserved powers brought into play by emergency conditions (pp. 473–74):

> The question is not whether an emergency furnishes the occasion for the exercise of that state power, but whether an emergency furnishes an occasion for the relaxation of the restrictions upon the power imposed by the contract impairment clause; and the difficulty



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> is that the contract impairment clause forbids state action under any circumstances, if it have the effect of impairing the obligation of contracts. . . . The Minnesota statute either impairs the obligation of contracts or it does not. If it does not, the occasion to which it relates becomes immaterial, since then the passage of the statute is the exercise of a normal, unrestricted, state power and requires no special occasion to render it effective. If it does, the emergency no more furnishes a proper occasion for its exercise than if the emergency were non-existent. And so, while, in form, the suggested distinction seems to put us forward in a straight line, in reality it simply carries us back in a circle, like bewildered travelers lost in a wood.

The dissenting opinion concluded: "If the provisions of the Constitution be not upheld when they pinch as well as when they comfort, they may as well be abandoned." The dissenters feared that the Court, to facilitate the extraordinary governmental measures being implemented at all levels, had begun to abandon the Constitution by embracing a reinterpretation so sweeping as to effectively destroy the traditional understanding and force of the constitutional protection of private rights.

During the Korean War, the Court had another occasion to consider its doctrine regarding emergency powers. President Truman, placed in an uncomfortable political position by a deadlocked union-management dispute that threatened a nationwide strike, directed the Secretary of Commerce in April 1952 to seize and operate several steel mills. The owners obtained an injunction to prevent the seizure. The Supreme Court upheld the injunction by a 6-3 vote in the *Youngstown* case decided June 2, 1952 (343 U.S. 579).

Although the Court denied Truman—by that time a very unpopular President—a power previously exercised freely by Wilson and FDR, the decision did not impose a restriction on the *government's* power to take private property in an emergency; it restrained only a *presidential* taking without specific statutory authorization (pp. 587, 653, 700). As Justice Black said in announcing the majority's opinion, "This is a job for the Nation's lawmakers, not for its military authorities." In a concurring opinion, Justice Jackson emphasized "the ease, expedition and safety with which Congress can grant and has granted large emergency powers." Neither the majority nor the minority gave any weight to private property rights. While the majority objected only to Truman's presidential high-handedness, the minority grumbled that "such a [presidential] power of seizure has been accepted throughout our history." The Constitution was read in this case, as in many orders, not as a bulwark against governmental oppression of private citizens but as the institutional setting within



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

which high officials in the different branches of government conduct their internecine struggles for supremacy.

By enacting the NEA and the IEEPA, Congress sought to institutionalize its regulation of the presidential exercise of emergency powers, but the Supreme Court has recently breathed new life into those powers, countermanding restraints codified in the 1976 and 1977 statutes. In 1983, the Court ruled in *Immigration and Naturalization Service v. Chadha* (103 *S. Ct.* 2764) that Congress may not use a one-house resolution *or any other method short of the full constitutional requirements for enactment of legislation* to overturn a decision delegated to the President. By implication, the ruling invalidated the use of a concurrent resolution, as provided in the NEA, as a restraint on the presidential exercise of emergency authority. Congress might have foreseen this outcome. In hearings on national emergency legislation in 1977, members of Congress were warned by administration witnesses and legal scholars that the proposed use of the concurrent resolution was constitutionally vulnerable. Against this background, skeptics might argue that Congress adopted the provision in order to achieve the appearance rather than the reality of congressional control. Whatever the congressional intention may have been, in *Chadha* one potential counterweight to the President's use of emergency powers was demolished by the judicial interpretation of the doctrine of separation of powers.

More significantly, the Supreme Court's two recent decisions directly involving the IEEPA all but gutted certain congressional constraints on presidential power set forth in the NEA. In *Dames & Moore v. Regan* (101 *S. Ct.* 2972 [1981]), the Court gave broad construction to the President's power to act under the IEEPA, endorsing President Carter's use of that act first to block Iranian funds and later to compel their return to Iran (thereby nullifying certain attachments issued by U.S. courts) as part of his effort to secure the release of American hostages. Noting other disputes over Carter's attempt to nullify Americans' legal claims to Iranian assets, the Supreme Court quoted approvingly a lower court's ruling in upholding these powers that "the language of IEEPA is sweeping and unqualified."[8] The Supreme Court (101 *S. Ct.* at 2983) ruled that the President did have the power to nullify the judicial attachments of Iranian assets secured by the American petitioners: "We . . . refuse to read out of Section 1702 [of the IEEPA] all meaning to the words 'transfer,' 'compel,' or 'nullify.' . . . both the legislative history and cases interpreting the TWEA fully

[8]Chas. T. Main Int'l. Inc. v. Khuzestan Water & Power Authority, 651 F. 2d 800 (1st Cir., 1981) at 806–807.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

sustain the broad authority of the Executive when acting under this congressional grant of power."

Equally significant was the Court's ruling on the scope of executive power to deal with such foreign policy disputes in the *absence* of statutory authority. While holding that President Reagan lacked statutory authority under the IEEPA to "suspend" U.S. citizens' claims against Iran pursuant to President Carter's agreement with the Iranian government, the Court nevertheless ruled that the President did not lack *constitutional* power to terminate the claims. Citing with approval Justice Frankfurter's statement in *Youngstown* that "a systematic, unbroken, executive practice, long pursued to the knowledge of the Congress and never before questioned . . . may be treated as a gloss on 'Executive Power,' " the Court held that "where, as here, the settlement of claims has been determined to be a necessary incident to the resolution of a major foreign policy dispute between our country and another, and where, as here, we can conclude that Congress acquiesced in the President's action, we are not prepared to say that the President lacks the power to settle such claims" (pp. 2990–91).

Lest anyone think that the judges did not understand the economic consequences of their decision, the Court rationalized its invalidation of the property rights involved by referring to the Iran-United States Claims Tribunal as an alternative forum established to provide relief in these cases. While unabashedly admitting that the American claimants had scant chance of success before the Claims Tribunal, the justices offered the consolation that the deprived parties "may well recover something on their claims" (p. 2990).

Further erosion of the NEA and the IEEPA came in *Regan v. Wald* (104 *S. Ct.* 3026), decided by the Supreme Court on June 28, 1984. At issue was the scope of the "grandfather" provision of the IEEPA, which exempts "existing uses" of Section 5(b) powers from the NEA's restrictions. The dispute arose over use of the Cuban Assets Control Regulations, first promulgated under the old TWEA, to prohibit Americans from engaging in economic transactions related to tourist and business travel to Cuba. At the time of the IEEPA's enactment, such transactions were permitted under a general license issued by the Treasury Department. In 1982, without a presidential declaration of national emergency or executive compliance with the procedures of the NEA and the IEEPA, the Treasury Department changed the regulations to prohibit most of the transactions. The issue before the Court was whether or not the dramatic change in the regulations was covered by the grandfather provision of the IEEPA and, therefore, exempt from the NEA procedures.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The issue had been discussed in congressional hearings on the IEEPA. The original grandfather proposal contained language that would have grandfathered not only existing uses of statutory authority but also any other authority exercised to deal with the same set of circumstances. Representative Bingham balked at such a broad formulation; he wondered "why it should be necessary to give [the President] authority to expand what has already been done." His understanding was that "grandfathering applies to *what has been done to date,* and that should be ample authority." The subcommittee's staff director, Roger Majak, responded that "it boils down to a question of whether we are grandfathering a particular situation, and all the powers that may be necessary to deal with the situation, or whether we are grandfathering the particular authorities themselves and their usage" (U.S. House of Representatives 1977b, p. 167, emphasis added). The upshot of the discussion was that the language of the grandfather provision was changed, eliminating any reference to the use of other authorities relevant to existing situations. A stripped-down clause emerged, grandfathering only "existing uses" of Section 5(b) authorities.

When the bill was presented for markup before the full House Committee on International Relations, Bingham stated unequivocally: ". . . we have grandfathered in *those actions currently being taken* under the Trading With the Enemy Act" (U.S. House of Representatives 1977a, p. 2, emphasis added). Under questioning by Representative Cavanaugh, Assistant Treasury Secretary Bergsten, the administration's spokesman for the IEEPA, reiterated the narrow purpose of the revised grandfather provision (p. 21):

> Mr. CAVANAUGH. First of all, Mr. Bergsten, would it be your understanding that [the grandfather clause] would strictly limit and restrict the grandfathering of powers currently being exercised under 5(b) to those specific uses of the authorities granted in 5(b) being employed as of June 1, 1977.
>
> Mr. BERGSTEN. Yes, sir.
>
> Mr. CAVANAUGH. And it would preclude the expansion by the President of the authorities that might be included in 5(b) but are not being employed as of June 1, 1977.
>
> Mr. BERGSTEN. That is right.

Notwithstanding this unmistakably explicit legislative history, the Supreme Court ruled by a vote of 5-4 that the grandfather provision *does* encompass changes in existing restrictions if they pertain to the same general "authority" previously exercised under the TWEA. The four dissenting justices (104 *S. Ct.* at 3046) could only observe that



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

"there is nothing in the language of the statute to suggest . . . that Congress intended the grandfather clause to provide a President with the authority to increase the restrictions applicable to a particular country without following the IEEPA procedures."

The immediate result of the Court's decision was that a major new curtailment of private travel to Cuba was implemented under the Cuban Assets Control Regulations without a declaration of national emergency or compliance with NEA procedures. With the stroke of a judicial pen, an apparent pinprick in the statutory armor against the abuse of emergency powers became a gaping hole.

## Conclusion

The history of the United States in the 20th century provides strong evidence that derogations from private property rights in a liberal democracy occur chiefly during national emergencies and that, once curtailed, private rights seldom regain their previous scope. The pattern should not be surprising. Crisis clearly alters the expected benefits and costs of curtailment of private rights on both sides of the political equation. A fearful public, ideologically predisposed to believe in the efficacy of governmental action, insists that the government "do something" to diminish the threat, perceiving the benefits of such action to be immediate and direct and the costs to be remote and largely external. This public perception is nurtured by those who, for material or ideological reasons, would use the occasion to further their economic or political aims. From a cost-benefit perspective, governmental officials experience reduced political costs and increased political benefits from curtailing private rights in crisis as compared to non-crisis conditions.

As people adjust to the crisis-expanded role of government, many variables change in ways that diminish the likelihood that the post-crisis retrenchment of the government will restore private rights to their previous scope. Private citizens discover that governmental action in a liberal democracy need not lead to the establishment of totalitarianism, as some conservatives have predicted. Governmental officials develop the bureaucratic technology to administer their controls less abrasively and more effectively. Many people find the politically and economically rewarding paths to personal advancement unique to systems powered by discretionary governmental authority. Thus, history is irrevocably altered by the crisis-induced expansion of governmental authority. The change is consolidated and compounded as new generations never experience the broader realm of private rights that once prevailed. For the younger genera-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tions, the status quo is the current, high degree of governmental power; for them there is no personal experience and, therefore, no genuine appreciation of the old regime.

The legal legacies of crisis tilt the polity in the same direction: Statutes, regulations, and judicial decisions expressing and facilitating expanded governmental powers become embedded in the law as well as in the public's consciousness. The plethora of crisis-engendered statutes and judicial decisions attests to the magnitude of this aspect of the politico-economic dynamics. The discretionary nature of governmental powers thereby created provides an easy avenue to their expanded use in future situations that the public perceives, or can be induced to perceive, as "crises."

Thus, private property rights, historically truncated during national emergencies, remain vulnerable to further erosion during future crises. Attempts to restrain the abuse of emergency powers have not eliminated the ratcheting effect of actual or purported emergency in augmenting governmental power. Only the respite of non-crisis affords time to contemplate and forestall the threat to liberty and private property rights inherent in the emergency psychology of the public and its exploitation by governmental officials.

## References

Alston, Lee J. "Farm Foreclosure Moratorium Legislation: A Lesson from the Past." *American Economic Review* 74 (June 1984): 445–57.

Baruch, Bernard. *Baruch: The Public Years*. New York: Holt, Rinehart and Winston, 1960.

Belknap, Michal R. "The New Deal and the Emergency Powers Doctrine." *Texas Law Review* 62 (1983): 67–109.

Bryce, James. *The American Commonwealth*, vol. 2. 3rd ed. New York: Macmillan, 1895.

Clark, Jane Perry. "Emergencies and the Law." *Political Science Quarterly* 49 (June 1934): 268–83.

Corwin, Edward S. *Total War and the Constitution*. New York: Knopf, 1947.

Gillam, Richard. "The Peacetime Draft: Voluntarism to Coercion." *Yale Review* 57 (June 1968): 495–517.

Greenberger, Robert S. "Reagan Plans More Sanctions On Libya Trade. Measure Would Ban Exports of Some Parts, Imports of Refined Oil Products." *Wall Street Journal*, May 28, 1986.

Higgs, Robert. "Crisis, Bigger Government, and Ideological Change: Two Hypotheses on the Ratchet Phenomenon." *Explorations in Economic History* 22 (January 1985): 1–28.

Higgs, Robert. *Crisis and Leviathan: Critical Episodes in the Growth of American Government*. New York: Oxford University Press, 1987.

Hoover, Calvin B. *The Economy, Liberty and the State*. New York: Twentieth Century Fund, 1959.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hughes, Jonathan R. T. *The Governmental Habit: Economic Controls from Colonial Times to the Present.* New York: Basic Books, 1977.

Karlin, Norman. "Substantive Due Process: A Doctrine for Regulatory Control." In *Rights and Regulation: Ethical, Political, and Economic Issues,* pp. 43–70. Edited by Tibor R. Machan and M. Bruce Johnson. Cambridge, Mass.: Ballinger, 1983.

Kempe, Frederick. "U.S. Officials May Loosen Libya Sanctions; Object Is to Prevent Qadhafi From Reaping $1 Billion Windfall From Firms." *Wall Street Journal,* February 3, 1986a.

Kempe, Frederick. "U.S. Plans New Sanctions Against Libya But Hopes to Avoid a Dispute with Allies." *Wall Street Journal,* August 26, 1986b.

McAdoo, William Gibbs. *Crowded Years.* Boston: Houghton Mifflin, 1931.

Murphy, Paul L. *The Constitution in Crisis Times, 1918–1969.* New York: Harper Torchbooks, 1972.

Navarro, Peter. *The Policy Game: How Special Interests and Ideologues Are Stealing America.* New York: Wiley, 1984.

Rossiter, Clinton. *Constitutional Dictatorship: Crisis Government in the Modern Democracies.* Princeton, N.J.: Princeton University Press, 1948.

Rossiter, Clinton, and Longaker, Richard P. *The Supreme Court and the Commander in Chief.* Ithaca, N.Y.: Cornell University Press, 1976.

Siegan, Bernard H. *Economic Liberties and the Constitution.* Chicago: University of Chicago Press, 1980.

Sumner, William G. *Essays of William Graham Sumner,* vol. 2. Edited by A. G. Keller and M. R. Davie. New Haven: Yale University Press, 1934.

Twight, Charlotte. "Government Manipulation of Constitutional-Level Transaction Costs: An Economic Theory and Its Application to Off-Budget Expenditure Through the Federal Financing Bank." Ph.D. diss., University of Washington, Seattle, 1983.

Twight, Charlotte. "U.S. Regulation of International Trade: A Retrospective Inquiry." In *Emergence of the Modern Political Economy,* pp. 139–95. Edited by Robert Higgs. Greenwich, Conn.: JAI Press, 1985.

U.S. Bureau of the Census. *Historical Statistics of the United States. Colonial Times to 1970.* Washington, D.C.: Government Printing Office, 1975.

U.S. House of Representatives. Committee on International Relations. *Revision of Trading with the Enemy Act.* Markup. 95th Congress, 1st Sess., 1977a.

U.S. House of Representatives. Committee on International Relations, Subcommittee on International Economic Policy and Trade. *Emergency Controls on International Economic Transactions and Markup of Trading with the Enemy Reform Legislation.* Hearings on H.R. 1560 and H.R. 2382. 95th Cong., 1st Sess., 1977b.

U.S. Senate. Committee on Banking, Housing, and Urban Affairs, Subcommittee on International Finance. *Amending the Trading with the Enemy Act.* Hearing on H.R. 7738. 95th Cong., 1st Sess., September 8, 1977.

U.S. Senate. Special Committee on the Termination of the National Emergency. *Emergency Powers Statutes: Provisions of Federal Law Now in Effect Delegating to the Executive Extraordinary Authority in Time of National Emergency.* Senate Report 93-549. 93rd Cong., 1st Sess., 1973.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Environmental Risk Management Through Insurance

*Martin T. Katzman*

## Introduction

Reported threats to human health from chemical releases into the environment have become a weekly ritual. Despite contrary biomedical evidence (Ames 1983; Peto 1984), a majority of the public remains convinced that chemical effluents in the environment are a major cause of cancer, cardiovascular disease, and birth defects (Zenter 1979). In response to concern with mass exposure to toxic chemicals from waste sites, the Resource Conservation and Recovery Act (RCRA)[1] and the Comprehensive Environmental Response, Compensation, and Liability Act[2] ("Superfund") were passed.

These two acts rely more on market incentives and less on traditional proscriptive or prescriptive strictures than any other environmental statutes. They require handlers of hazardous chemicals to establish "financial responsibility," which means a guarantee to pay for damages up to a specified limit. Unless a firm can meet the test for self-insurability, financial responsibility must be met by insurance (U.S. EPA 1982). Through a demand-pull strategy, the acts

---

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics and Environmental Sciences at the University of Texas at Dallas and Senior Economist at Argonne National Laboratory. This project was supported with grants from the Huebner Foundation for Insurance Education and Resources for the Future. Earlier versions of this paper were presented to the Risk Theory Seminar, Nashville, April 1985 and the Seminar on Risk Management: Through Regulation or Insurance?, Cato Institute, May 1985. Comments from the following individuals are gratefully acknowledged: J. David Cummins, Robert C. Witt, Jr., Anthony M. Champagne, Murray Leaf, Joe G. Moore, Jr., Karl Borch, Teivo Pentikainen, Howard Kunreuther, Emilio Venezian, Catherine England, and Fred L. Smith, Jr. The usual caveat applies.

[1]42 U.S.C. 6924 (a)(t); 40 C.F.R. 264.14 and 265.14 et seq.

[2]42 U.S.C. 9608 (1986).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

encourage the establishment of a market in insurance for nonsudden or gradual pollution damage.

When RCRA was enacted in 1976, several London insurers had been developing liability policies for nonsudden or gradual pollution accidents. When Superfund was passed in 1980, a few American insurers were offering such policies. By 1983, despite the misgivings of many underwriters (U.S. Treasury 1982), at least a dozen primary insurers were offering coverage. In addition, more than 40 insurers had established a reinsurance pool to spread the risks further (Katzman 1985, chap. 5). The market-based approach to chemical risk management appeared to be on the road to success.

By the end of 1984, the pollution-insurance initiative lay in shambles. London reinsurers had withdrawn from the market, carrying many existing and prospective insurers in their wake. Pollution insurers numbered about eight worldwide, most of whom insured small-scale facilities like gas stations and dry cleaning establishments. Only two insured "heavy" risks, like chemical plants, from which catastrophes are most likely to result. By the summer of 1986, only one insurer offered coverage for gradual pollution occurrences.[3]

While there have been some signs of recovery, the pollution liability market has not developed as rapidly as the federal government had hoped. As a result, Congress twice had to postpone implementation of the financial responsibility requirements of RCRA. The development of financial responsibility requirements for generators under Superfund has been aborted.

Does the collapse of the market indicate a fundamental problem of insurability, a remediable imperfection, or merely a cyclical disequilibrium? This query raises broader questions about: (1) public policy alternatives for the management of catastrophic risks; (2) the inherent insurability of pollution liability; (3) the conditions necessary to sustain a viable pollution liability market; and (4) the value of insurance as a mechanism for rendering private risk-management decisions socially acceptable.

The fact that the insurance industry (Cheek 1982) perceives its mission as risk-spreading rather than risk-eliminating is irrelevant to the inquiry. In political economy, the distinction between the conscious purpose of actors and their social function is as venerable as the concept of the invisible hand.

[3]"Difficulty in Obtaining Liability Insurance Said Major Problem for Waste Facilities," *Environmental Reporter* 15 (15 February 1985): 1660–61.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

## Public Risk Management Policy

Advanced industrial economies both eliminate and create risks. The whooping cough vaccine that saves millions from infectious disease may cause brain damage to an unlucky few. The insecticide that saves masses from starvation may explode during its production, as in Bhopal. Commercial enterprises internalize some of the social benefits of these risk-reducing activities through higher profits, while most are enjoyed by the public. The question of who bears the social costs remains open.

The social costs of the risk-creating activities are readily internalized by the firm when the risks are knowingly and voluntarily taken by a potentially injured party. Where voluntary exchange exists, social costs can be signaled through labor or product markets. As implied by the Coase theorem, bargaining between workers and their employer can result in a mutually acceptable risk management package that includes a wage premium, worker compensation in the case of accidents, and some residual or retained risk. For example, workers demand higher wages for more dangerous jobs, thereby inducing employers to undertake marginal safety expenditures that cost less than the expected marginal compensation payments (Viscusi 1979). Similarly, consumers may be willing to pay more for a safer tool, to the greater profit of the manufacturer.

Where accidents affect third parties, the achievement of mutually acceptable risk management is not automatic. Potential victims of environmental pollution are not always knowledgeable about subtle exposures, such as carcinogens in the water supply. Even if they were, potential victims cannot easily signal their distress by market exchange. The transactions cost of arranging Coasean contracts between a multitude of dispersed, potential victims and one or more polluters may be prohibitive.

There are three major regimes for forcing the responsible actors to internalize the costs of accidents to third parties: tort law, statutory regulation, and user charges (Shavell 1984; Landes and Posner 1984). Activated subsequent to an accident, the tort process determines which party, if any, is liable for the ensuing costs. The anticipation of potential liability in the future creates a powerful influence on current behavior.

Statutory regulation attempts to constrain behavior prior to the occurrence of an accident. Regulations may mandate certain inputs (for example, clay liners in hazardous waste sites). Other regulations may constrain effluent volume or completely prohibit specific outputs (by banning the production of a chemical). Statutory sanctions



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

may include civil and criminal penalties such as fines, injunctions, and imprisonment.

User charges are fees imposed on behavior likely to result in external diseconomies, as it occurs. User charges share characteristics of both statutory regulation and the tort process. Like regulation, a user charge is implemented by state action; like the tort process, a user charge is a market mechanism that leaves risk management decisions in private hands. Traditionally, tort recoveries are assessed only if harm results and fault is determined, while user charges are assessed whether or not harm results, without regard to fault. Through the tort process, a victim can recover damages, but he does not automatically recover through a user charge regime.

The three regimes for controlling accidental external diseconomies differ in their efficiency, distributional effects, and transactions costs (Calabresi 1970). The efficiency criterion weighs the expected marginal cost of pollution against the marginal cost of pollution control. Where these marginals are equal, the cost of pollution-engendered accidents plus the cost of accident prevention is minimized.

Third-party risks raise equity issues that are more compelling than first- and second-party risks. When the risk is borne by the beneficiary of the risky activity, such as the chain smoker, no blatant inequity calls for publicly sanctioned redress in the event of loss. When a third party becomes ill because of contamination of a water supply, the assumption of risk is neither knowing nor willful.

Equity demands symmetry between the costs and benefits of risk bearing. In the case of an accident, this means spreading the costs from the few who bear the brunt of the losses to a broader class of beneficiaries. In the environmental context, the equity criterion is only partially captured by the slogan "the polluter pays." Equity also demands the completion of the financial transfer, that is, victim compensation.

While the efficiency and equity objectives are analytically distinct, the public finds least acceptable those risky activities where the distribution of benefits and potential losses is asymmetric (Slovic, Fischhoff, and Lichtenstein 1980). This suggests that a tort regime, which facilitates compensation of potential victims by the injurer, can render a given hazard more acceptable than a regulatory regime, which does not provide for compensation.

All societal risk-management regimes incur substantial transactions costs. These costs include the overhead of establishing rules, monitoring behavior, analyzing risks, enforcing rules, and spreading or readjusting losses when accidents occur. These transactions costs



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

are deadweight drains from the gross efficiency gains of internalization (McKean 1980; Kakalik et al. 1984).

Efficiency and equity do not necessarily exhaust all of the attributes of risk acceptability. Libertarian values are clearly a component of risk acceptability (Abraham 1986, chap. 2). The public may reject a hazard that met both efficiency and equity criteria because the risk would be imposed rather than voluntary. Indeed, a voluntarily chosen technology may be more acceptable than an imposed technology that results in a hundredfold fewer fatalities (Slovic, Fischhoff, and Lichtenstein 1980).

As a regime that internalizes risks after an accident, a tort system by itself faces inherent limitations. In some instances, a polluting firm was dissolved before damages were discovered. In others, damages exceeded the firm's net worth (Katzman 1985, pp. 67–68). The Supreme Court has ruled that a functioning industrial polluter can escape an order to clean up a toxic waste site under the umbrella of federal bankruptcy.[4] Obviously, a business has no incentive to reduce the probability of accidents whose consequences become visible after dissolution or for which losses exceed net worth.

In an attempt to block an escape from liability through dissolution or bankruptcy, both RCRA and Superfund established financial responsibility requirements. Under RCRA, facilities that treat, store, or dispose of hazardous chemicals must prove financial responsibility for sudden accidents, at the levels of $1 million per occurrence and $3 million annually, and for nonsudden occurrences, at levels of $3 million and $6 million, respectively. Under Superfund, additional financial responsibility requirements were to be established by the end of the 1980s for generators. These requirements can serve as the basis of a user-charge regime.

Economists generally favor user charges as the most efficient regime for internalizing the third-party cost of pollution (Milliman 1982). For routine effluents, the charge is assessed on the basis of volume, such as pounds of sulfur dioxide. This charge becomes a cost of production that is passed on to the consumer.

An effluent fee on low probability–high severity chemical hazards is unworkable because the accident activating the user charge may never occur, despite the ever-present risk. Moreover, charging for gradual releases that remain hidden, before surfacing years later, is infeasible. Finally, the charge could be avoided by bankruptcy.

The 1980 Superfund act introduced two user charges. First, a production and import tax on 40 specific chemical compounds was ear-

[4]Ohio v. Kovacs, U.S.S.C. 83-1020.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

marked for an emergency cleanup fund. This tax, whose rate is proportional to the presumed riskiness of the substance, is a poorly honed market mechanism; it is insensitive to the differences between careful and careless chemical handling.

A second tax, levied on hazardous waste, was earmarked for monitoring retired waste sites that have come under federal jurisdiction. This tax indirectly encouraged the reduction of waste flows, which correlate with, but are not equivalent to, the hazard. The repeal of this tax in the 1986 amendments (Title V) and the imposition of a broadly based corporate income tax are moves away from the virtues of user charges.

*The Potential of Insurance*

Liability insurance blends the advantages of both the tort system and user charges as instruments for internalizing the costs of accidents. Under the following conditions, insurance encourages "acceptable" risk management practices, with minimal statutory regulation:

1. Tort rules result in predictable assignments of liability for accidents, based upon reasonable judgments about causal likelihood and relative responsibility. While these rules may evolve in slightly unanticipated ways, the insurers, the insured, and potential victims can rely upon them in managing risks. These rules do not pose impenetrable barriers to recovery to plaintiffs.
2. The courts respect the sanctity of the insurance contract. In other words, both the "reasonable expectations of the insured" and the insurer's "reliance interest" are honored (Abraham 1986, chap. 5).
3. By the application of its skills in safety engineering and actuarial science, the insurance industry can identify and assess the risks of alternative chemical products and processes.
4. Insurers set premiums on the basis of expected losses, which are sensitive to safety measures taken by the chemical industry.
5. Competitive pressures among insurers result in continual improvements in the art of risk analysis. Unlike government bureaucrats, insurers suffer penalties—that is, they lose business—by overestimating risk and setting premiums too high. If they underestimate risks, they suffer financial losses.
6. With the help of insurers and risk analysts, corporate risk managers select cost-effective products and processes, which minimize the sum of insurance premiums, expected payments to victims (in excess of insurance coverage), and risk-reduction expenditures (Ehrlich and Becker 1972).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

7. Because the costs of accidents are internalized and polluters have the financial means to pay victims, both efficient deterrence and just compensation result from private decisions. The public will therefore find private risk management acceptable.

The idealized role of insurance as a regulatory tool is most closely fulfilled in the fire insurance industry. Here insurers have developed a tradition of research into fire safety and an aggressive search for cost-effective risk-reduction practices, inspection programs, and merit rating (Denenberg et al. 1974, pp. 82–85). Fire losses would undoubtedly be greater without these insurance industry activities.

Behavioral decision theory casts doubt on whether firms would indeed act as hypothesized by normative insurance theory for gradual pollution exposures. According to this theory, risk managers have a "finite reservoir of concern," and they pay little attention to low-probability events (Slovic et al. 1977; Schoemaker and Kunreuther 1979). Because of turnover, risk managers may reap no reward within the organization for reducing remote, future risks.

Agency theory, however, suggests that the insurance industry performs the function of protecting the stockholders' long-term interests. The underwriting process serves as a control mechanism that requires the formal justification of risk-management decisions (Mayers and Smith 1982). Financial responsibility requirements are an additional control mechanism that focuses the attention of managers on rare chemical accidents. Preliminary experiments with corporate risk managers suggest that they indeed pay attention to the low-probability, high-consequence contingencies as suggested by normative insurance theory (Katzman 1986). These considerations suggest that demand should not limit the development of a pollution liability insurance market. The uncertainties lie on the supply side.

## Insurability of Chemical Risks

### *Conventional Standards of Insurability*

The enactment of a financial responsibility requirement for nonsudden liabilities does not automatically create a market in pollution liability insurance. Although someone will insure any risk at a high enough price, all risks are not readily insurable at prices that will not drive the beneficial, risk-creating activity from the market.

Readily insurable exposures have several well defined characteristics (Denenberg et al. 1974, pp. 154–56; Mehr and Cammack 1980, pp. 25–35). First, exposures must be homogeneous, numerous, and uncorrelated enough to allow risk pooling. Second, the insured must have no incentive to bring about the loss. Third, the fact of the loss



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

must be clearly determinable. Finally, the loss must be frequent enough to calculate pure premiums. How close do chemical hazards come to these ideals?

*Number and Homogeneity of Exposures*

The number of potential exposures is relatively large. In the United States there are 115,000 chemical plans, 5,000 transporters, and at least 100,000 industrial waste disposal sites, of which at least 30,000 contain hazardous wastes (U.S. Congress 1980, chap. 2; Katzman 1985, p. 77). With the diffusion of industry standards of care, all of these facilities should become more homogeneous.

Most firms that meet the RCRA financial test do self-insure. Of those that purchased insurance in 1984, few obtained more than the statutory minimum (Katzman 1985, p. 93). Nevertheless, demand is likely to grow swiftly, as generators come under the financial responsibility requirements of Superfund. Ultimately, the number of exposures should be sufficient to achieve adequate risk pooling.

A major source of correlation of losses is the generic nature of many chemical products. If a given chemical is discovered to cause latent environmental harm, all manufacturers might be held jointly and severally liable. If an insurer concentrated its portfolio only upon generators or handlers of that chemical, then it would face the same correlated risk as an insurer whose portfolio consisted of hurricane insurance on the Gulf Coast. An underwriter could avoid this eventuality by insuring across many chemical products.

*Perverse Incentives to Bring about the Loss*

In the economic approach to the law, all accident-engendering behavior is subject to the control of the firm (Ehrlich and Becker 1972; Landes and Posner 1984). In deciding to produce a particular commodity with a particular technology, the firm chooses a given level of risk. The decision to undertake a risky activity, however, is not the same as the choice to cause an accident. In this respect, chemical exposures are no different from common mechanical exposures. As discussed below, the doctrine of joint and several liability in toxic torts reduces incentives to choose safe technologies.

*Definiteness of Loss*

Losses that cannot be publicly verified lend themselves to counterfeit claims. Property damage, ecosystem contamination, and many personal injuries such as tumors or birth defects can be publicly validated. Until recently, public policy toward hazardous chemicals has focused exclusively upon such injuries.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Recent research, however, questions the use of cancer as the epitome of environmental disease. Continued exposure to chemicals results in increased sensitization, that is, a lowered threshold of morbidity response. Moreover, morbidity also may take the form of a diffuse malaise, analogous to the debilitation associated with lead poisoning (Katzen 1985). In *Ayers v. Jackson Township*,[5] plaintiffs argued that a contaminated municipal water supply was responsible for malaise, rashes, and general anxiety. While the trial court rejected such reports of malaise or rashes as evidence of injury, it awarded the plaintiffs $2 million for emotional stress and $8.2 million for lifetime medical monitoring. An appeal overturned the recovery for stress and medical monitoring, and the case is now at the portals of the New Jersey Supreme Court. Regardless of this outcome, there is no guarantee that future courts will not recognize malaise as a compensable ailment in its own right. Indeed, in *Jackson v. Johns-Manville*, a federal court of appeals upheld an award for anxiety over probable future illness.[6]

*Calculability of Loss Distributions*

Because chemical disasters are such rare events, the computation of premiums on the basis of loss experience may be virtually impossible. Even if historical loss data were available, they would reflect outmoded safety technologies. But this problem is characteristic of any complex, innovative, technological system, like a satellite or offshore oil rig, and is not unique to chemical catastrophes.

Underwriters claim little knowledge of the risk-reduction technologies available to chemical handlers, and they express little interest in acquiring such knowledge (Cheek 1982). They believe that developing expertise would be costly and argue that the insured would conceal proprietary, technical information (Katzman 1985, pp. 86–87; AIRAC 1985).

The lack of expertise has proven to be a bugaboo. The profit potential in pollution liability insurance has given birth to firms that specialize in risk analysis. While their art is relatively immature, environmental risk analysts have developed workable techniques. Simple screening methods have been employed as a first cut (U.S. EPA 1977; Harris et al. 1982). These techniques score facilities on the basis of such risk factors as the nature of chemicals handled and proximity to populations.

[5]N.J. Super.L., 461 A.2d 184, 189 N.J. Super.561 (1983), 493 A.2d 1314 (1985).
[6]54 U.S.L.W. 2400 (5th Cir. 1986).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The quantity and quality of risk analysis is likely to grow with the size of the pollution liability market. Eventually, underwriters may adopt more sophisticated methods, like event-tree and fault-tree analysis (Katzman 1985, chap. 6). Developed for weapons systems and nuclear power plants, these techniques trace the chain of events that could unleash a catastrophe.

Technological risk analysts tend to be overconfident of their ability to assess low probabilities and often demonstrate upward biases in their estimates of the reliability of complex technological systems (Fischhoff, Slovic, and Lichtenstein 1978). Insurers tend to be more cautious than risk analysts and set premiums accordingly. For example, the near-meltdown at Three Mile Island is viewed as far more likely under the assumptions of the insurance industry than under those of the technological risk analysts and regulators.[7] Nevertheless, underwriters can view risk-analytic results as a lower bound of expected losses.

## Chemical Damages and the Tort Process

The fixing of insurance premiums depends critically upon the expected tort claims against the insured. The insurer must predict not only the frequency and severity of accidents, but also the distribution of tort claims against the insured. Unfortunately, claims settlement for chemical accidents is far less predictable than the settlement of claims for commonplace mechanical accidents (Best and Collins 1982).

First, many toxic chemicals persist in the environment. An accidentally released chemical may gradually seep into the groundwater or may become concentrated as it accumulates in the food chain. As a result, the time of a spill or release may precede the time of human exposure by many years. Second, for some toxic chemicals, particularly carcinogens, the time of human exposure may precede the manifestation of injury by several decades. Chemical exposures to one generation may even harm their offspring. Under traditional tort law, a statute of limitations prevents the filing of a suit more than three to five years after an accident. Because of the time lags between

[7]If the Bayesian prior probability of a meltdown is 1/30,000, as taken from a report by a distinguished nuclear engineer to the Nuclear Regulatory Commission, then the probability of one meltdown in 500 reactor-years is only 0.015. If the Bayesian prior probability is 1/1060, as revealed in premiums of the nuclear liability pool, then the probability of one meltdown in 500 reactor–years is 0.27. While Three Mile Island was not a meltdown, it came close and makes the insurers' prior probability appear more plausible than that of engineers and regulators. See Wood (1981).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

chemical release and human response, conventional statutes of limitations may prove an insuperable barrier to recovery.

Third, there are multiple pathways through which a given chemical may reach humans. These include the environment, the workplace, and the home. A given biological response, such as lung cancer, may result from one or several alternative chemicals. The multiplicity of potential pathways raises two problems in establishing liability: identification of the defendant and proof of causation.

Under traditional tort law, the plaintiff must identify one or more specific defendants. If an environmental release can be traced to action or inaction of one identifiable party, then specific liability can be assigned. In many cases of gradual environmental release, however, identifying that defendant whose molecule caused the plaintiff's exposure may be virtually impossible. At an abandoned hazardous waste site, for example, many companies may have discarded identical chemicals.

Proof of causation depends upon presenting sophisticated biomedical evidence, most of which is indirect or analogous in nature. Epidemiologists have difficulty sorting out health effects of life-style (especially diet and voluntary consumption of stimulants) as well as occupational and environmental exposures to chemicals. Toxicologists are uncertain about long-term human responses to low, intermittent doses of chemicals, especially when many chemicals act in concert. Not surprisingly, expert testimony about causality is rarely conclusive.

The latency period between exposure and manifestation of illness obfuscates the search for causality. The latency period increases opportunities for further confounding causes to intrude. Moreover, the quantity and quality of evidence (including eyewitnesses) decays over time.

Advances in biomedical measurement may make the assignment of liabilities even more difficult. As lower and lower concentrations of chemicals in the environment become measurable, the number of chemical hazards for which human exposure can be measured, and hence the number of alternative explanations for an injury increases. As techniques for diagnosing vague symptoms improve, the potential number of measurable adverse effects increases as well (Katzen 1985).

### *The New Toxic Torts*

The barriers of traditional tort law virtually prohibited the recovery of damages from long-term chemical injuries. Indeed, the inability of victims to internalize the costs of environmental damage through the tort process is a probable cause of the growth in environmental



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

regulation. In the past decade, changes in toxic tort law have been nothing less than revolutionary, although uneven from state to state. In some states barriers to recovery in toxic tort actions have also been reduced by statute (U.S. Congress 1980; Priest 1985).

Over 40 states have abandoned traditional statutes of limitations by adopting a "discovery rule." Under such a rule, the "clock begins to run" not with the chemical release but with the victim's exposure or the manifestation of injury. A few state legislatures, however, have reaffirmed the traditional statute of limitations, which begins at the time of the alleged exposure. In such states, recovery for long-latent damages from chemicals is nearly impossible, with exceptions for victims of Agent Orange and asbestos (Katzman 1985, chap. 3).

Landmark cases in the areas of pharmaceuticals and asbestos have addressed the problem of apportioning liability among defendants for a specific chemical exposure. Most courts have adopted rules of joint and several liability.[8] A chemical manufacturer may be joined as a defendant with a contractor (such as a transporter), a competitor who dumped a similar chemical in a common site, or a successive landowner. While a court may apportion liability according to the relative share of hazardous chemical produced or discarded, it need not do so. Once a joint defendant is found liable, the burden of apportionment is transferred from the plaintiff to the defendant. If other parties are unidentifiable or bankrupt, then one viable defendant may be held liable for all the damages. This rule of apportionment increases the probability that *someone* pays for environmental risks imposed on third parties.

While this rule serves the goal of compensating victims, it is counterproductive to efficient deterrence. By making a firm "its brother's keeper," the costs of poor waste-handling practices are spread to other firms, and incentives for deterrence are attenuated. Joinder attenuates incentives for risk reduction because the benefits may be shared by the industry rather than internalized in the firm (Katzman 1985, pp. 57–60). By employing joint and several liability to search for the "deepest pocket" for purposes of compensation, the courts have created a "moral hazard." The emergence of mutual institutions to offset the disincentives for deterrence is discussed below.

[8]Borel v. Fibreboard Paper Products Corp., 493 F.2d 1076 (5th Cir. 1974) permitted asbestos workers to sue the suppliers to their employers, without identifying one specific supplier whose asbestos caused disease. In Sindell v. Abbott Laboratories, 26 Cal.3d 588; Cal. 607 P.2d 924; 163 Cal.Rptr. 132 (1980), plaintiffs ingested a drug to suppress miscarriages. Several daughters developed cervical cancer during puberty. The court apportioned damages according to the manufacturers' market shares.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In tort actions involving hazardous chemicals, the courts have virtually abandoned the theory of negligence in favor of strict liability, as they had earlier in the areas of worker compensation and product liability (Priest 1985). While the plaintiff no longer has to prove that the defendant failed to exercise care in specific ways, recovery under strict liability still depends upon a formidable proof of scientific causation.

*Insurability Consequences of Legal Risks*

Underwriters have had difficulty in anticipating the behavior of the courts, and they seriously underestimated the legal risks of pollution exposures. Despite their caution, the premiums that insurers collected in 1983 were about one-third of the losses expected after eventual claims settlement.[9]

To offset their losses, it seems logical that underwriters would simply increase their premiums by a factor of three or more until they surpass the break-even point. There are two possible explanations why they have not done so: differential response of insurers and insured to ambiguity, and the adverse selection problem.

First, low probability losses are poorly calibrated and consequently ambiguous. Hogarth and Kunreuther (1985) have performed some interesting experiments on the impact of ambiguity about loss frequencies on both the offering price and selling price of insurance. They found that for very low-probability losses, insurers demand a premium that exceeds the expected losses by a substantial margin. The insured firms were willing to pay less than the expected loss, a common result (Slovic et al. 1977; Hershey and Schoemaker 1980). The more ambiguous the estimate of the frequency, the greater the divergence between required premiums and willingness to pay.

Second, in setting premiums, insurers may be less capable of discriminating among risk classes than the insured, an example of the "lemon" problem (Akerlof 1970). Premiums based on industry averages encourage firms that believe they are less risky than the average to self-insure. This exodus leaves the insurers with the riskier exposures. The ratio of losses to premiums then rises, the insurer raises its premiums, driving additional firms into self-insurance, ad infinitum. Insurers could attempt to reduce adverse selection by acquiring more information, but this is costly. Acquiring more information raises the overhead and premiums, further exacerbating adverse selection.

[9]"Insurers Warned to Write Pollution Coverage," *National Underwriter* (25 July 1985): 28.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Although the empirical relevance remains to be tested, these arguments suggest that an insurance market for low-probability risks might not emerge in the absence of financial responsibility requirements. Several states prohibit self-insurance and set the required insurance above the federal levels. These stricter requirements for financial responsibility do not appear to have affected the attractiveness of these states for either the insurer or the insured (AIRAC 1985, Table 5).

## The Rise and Fall of the Market

Several major oil spills in the 1960s raised public and industry awareness of a whole class of accidents that could have long-term consequences. In response, the basic commercial insurance policy—the Comprehensive General Liability (CGL) form—written after 1971 appended a clause that excluded nonsudden or gradual pollution "occurrences." These were defined as continuous or repeated releases of pollutants that resulted in unanticipated or unexpected damages to persons and property (Tyler and Wilcox 1981). The exclusion did not apply to sudden chemical accidents, such as explosions, which are similar in temporal demarcation to mechanical accidents.

This exclusion created a gap in coverage and a potential opportunity. In response, Environmental Impairment Liability (EIL) policies were developed on the London market by 1973. As they have evolved, EIL or pollution liability policies have been tailored to the distinction between sudden and nonsudden pollution (Katzman 1985, chap. 5).

For common mechanical accidents the dating of the accident and loss is not difficult. As noted, however, there may be significant time lags between the time of a chemical release, human exposure, and manifest injury. The difficulty of defining the time of a nonsudden chemical release obfuscates the activation of conventional insurance policies, which are triggered by an occurrence. The insured may have been covered by several underwriters during the relevant period. The losses cannot easily be allocated among a sequence of insurers, who can never be sure that the books are closed on any given policy year.

To obviate this difficulty, EIL policies are issued on a "claims-made" basis rather than an occurrence basis. The insured is covered for claims made during the policy year, even if the gradual or recurring chemical release or human exposure occurred prior to the policy date.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

EIL policies specify a retroactive date for such occurrences. At first glance, a retroactive date appears redundant because the policies are not activated by the pollution occurrence. Since businesses acquire and divest specific operations, it is important for the insurer to know how far back in time information must be collected. If an especially hazardous operation were divested before the retroactive date, current claims for resulting damages would not be honored.

While insurers have differed in their terms, all EIL policies cover liabilities for bodily injury and property damage to third parties, the cost of legal defense of a claim, and off-premises cleanup of a preventive nature. Some policies cover on-premises cleanup of a preventive nature, when there is an imminent off-premises hazard. All policies exclude damages resulting from willful violation of regulations and the costs of routine cleanup.

The intent of the insurance industry was to segment the market between liabilities from sudden occurrences and those that were gradual and nonsudden. The former, which includes conventional mechanical accidents, would continue to receive coverage under the traditional CGL policy, with its pollution exclusion. The latter, which includes latent pollutant damages, would be covered under the new EIL policy.

The CGL and EIL policies are quite different in their regulatory implications. The CGL policy is forward-looking, because incidents that occur in the future as a result of behavior in a given policy year are covered by that year's policy. The risk analyst and underwriter thus look at the future stream of losses resulting from this year's actions. Risk-based premiums thus have a deterrent effect on current risk-management decisions.

The EIL policy is backward-looking, because it is activated by claims made in the policy year resulting from past actions. The risk of insuring firms that were heavily engaged in handling chemicals (or shared facilities with such firms) depends more upon past risk-management practices than current ones. Chemicals released into the environment several decades ago may already have initiated latent diseases. While the past cannot be undone, some of the consequences can be.

In some cases, chemical spills are like a slow avalanche, which creates a present danger with ample warning. If firms are capable of remedial action, such as cleaning up older waste sites or purifying damaged aquifers, current behavior can affect current liability. Under the slow avalanche model, the current premium structure can reward mitigation efforts that reduce the probability and severity of an accident.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Nevertheless, the EIL policy may have perverse effects on deterrence. The claims-made format internalizes costs less efficiently than the occurrence format. Insurance premiums are factored into the cost of production, which is passed on to current consumers. Premiums on occurrence policies reflect future risks resulting from current production, and efficiency requires that current consumers pay these costs. Premiums on claims-made policies, in contrast, reflect risks from the past. Consumers who benefited from hazardous products in the past cannot be assessed once damage is discovered. The more involvement a firm had with hazardous chemicals in the past, the higher will be its claims-made premiums and the more its current customers will be overpaying. For new chemical firms, premiums will not reflect a legacy of the past, and its consumers will be underpaying for future risks (Abraham 1986, chap. 3).

Because they have no toxic skeletons in the closet, new firms may be able to purchase EIL insurance for lower premiums than established ones. This cost advantage might encourage the creation of new fly-by-night chemical or waste-management firms, which drive older firms out of business. This contingency raises the underwriters' risk, because Superfund authorizes direct claims against insurers of abandoned facilities.

In other respects, a claims-made policy is less risky for insurers than an occurrence-based one. In underwriting a claims-made policy, the insurer is not making a commitment into the indefinite future, when liability rules, medical detection technology, and jury awards may differ from today's. Furthermore, under a claims-made format, if an insured switches insurers from year to year, the liability falls unequivocally upon the insurer at the time the claim is made.

The claims-made format also imposes risks upon the insured. Suppose an insured firm undertakes a hazardous activity for a single year only, as a small part of a larger operation. This firm will have to purchase an EIL policy for the indefinite future, without any guarantee that coverage would be available in the future when damages may become manifest.

These considerations indicate great mutual advantages to long-term "monogamous" contracts between the insured and the insurer. Such a contract might specify annual adjustments for loss experience or for changes in interest rates. It might be voided if a facility were downgraded below industry standards. If the term of the contract is long enough, say more than 10 years, the distinction between claims-made and occurrence policies blurs.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

## Resuscitating the Market

Reductions in the number of pollution liability insurers are occurring just when federal and state financial responsibility requirements are being extended. As risk managers are demanding more insurance coverage, the supply is diminishing. How can this disequilibrium be resolved?

The collapse of the pollution liability insurance market in 1984 was partially a result of cyclical readjustments in reinsurance markets. The practice of setting premiums on the assumption of high-interest earnings proved disastrous when interest rates fell.[10] Unexpectedly high losses resulted from large mechanical accidents (such as satellite losses or oil rig collapses). Underwriters allege that large indemnity payments for toxic litigation contributed to their loss of underwriting capacity, although critics (Anderson 1985) contend that insuring asbestos manufacturers has been profitable.

To some extent the shortage of reinsurance is self-correcting, as premiums rise and new capital flows into this sector. The across-the-board spurt in insurance premiums and reduction in availability of coverage after 1985 reflect this equilibrating process. Market volatility also may be an aspect of learning how to insure a new line.

If the collapse of the pollution liability market were merely a cyclical or learning-curve effect, then the market would revive spontaneously. In this case, the proper public policy is to do nothing. So long as premiums are unregulated, the supply of insurance should be forthcoming.

There remain fundamental problems of insurability that result from legal risks. Judicial decisions in three areas have undermined the predictability of insurance contracts: (1) activation of liability policies for first-party damages; (2) reinterpretation of the pollution exclusion; and (3) adoption of theories of joint and several liability.

### *Indemnity for First-Party Cleanup*

A liability insurance policy is intended to be activated by damage to third parties, not to the insured itself. EIL policies are not intended to indemnify insured firms for on-site cleanup necessary to comply with the law. Nevertheless, the courts increasingly require insurers to indemnify polluters for cleaning up their own property under the RCRA provision of preventing an imminent hazard to third parties (Schmalz 1982; Aickin 1985).

[10]"Money Management by Insurance Industry Blamed for Lack of Environmental Coverage," *Environmental Reporter* 16 (24 January 1986): 1789–90.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Underwriters argue that in arbitrarily rewriting insurance contracts, the courts are using their industry to finance a social program. The insurer's pocket is not so deep as the public believes. The $100–$200 billion estimated cost of cleanup (Aickin 1985) exceeds the $48 billion surplus and the $90 billion premium income of American property-casualty insurers (Huebner, Black, and Cline 1982, pp. 514–15).

While compelling in some respects, the insurance industry's argument is flawed. Rulings to indemnify the insured for the costs of on-site cleanup are efficient if the cost is less than the expected cost of off-site damage, that is, the probability of damage multiplied by its severity. Insurers might have to pay even more if the damage were permitted to occur. Indeed, voluntary efforts provide some evidence that some cleanups cost less than expected damages (Clean Sites 1984). This does not imply, however, that all EPA-ordered cleanups are cost-effective.

In principle, there is no reason why all insurers could not knowingly underwrite policies on cleanup costs, as several now do (AIRAC 1985). Insurers routinely cover accidents that have already happened, gambling on collecting a premium greater than the discounted settlement costs (Smith and Witt 1985).

*Reinterpreting the Pollution Exclusion*

The courts increasingly ignore the pollution exclusion by redefining gradual pollution as "sudden" and "accidental" from the standpoint of the insured's knowledge and intent (Rich 1985). In essence, the courts have transformed the CGL policy into a pollution liability policy with unlimited coverage. Because of the gradual nature of pollution, the limits of previous years can be activated ad infinitum once the coverage of one year has been exhausted. While such a layering of policies has not occurred so far, there is no guarantee that it will not in the future (Anderson 1985).

Uncertainty about judicial interpretation of *past* insurance contracts does not render future environmental liabilities uninsurable. If underwriters refuse to insure a single exposure henceforth, they will still encounter liabilities under old CGL policies. To obviate future confusion, the industry's Insurance Services Office has tightened the pollution liability exclusion in the CGL form. As of January 1986, the form did not include coverage for sudden and accidental pollution. On a prospective basis, this exclusion neatly partitions the market for pollution-related accidents from other risks. Contractual confusion about whether a particular incident is sudden or gradual is irrelevant in a consolidated claims-made EIL policy.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*Joinder of Defendants*

The widespread application of joint and several liability in toxic torts remains the single most important obstacle to insurability. When defendants are joined, an underwriter faces the risk of indemnifying a client for claims resulting from damages caused by other firms. In setting premiums, the underwriter can analyze the inherent risks of its insured, but it cannot easily assess the risks resulting from joinder. Although the underwriter can offer incentives to its client for reducing risks, it has no contractual means of influencing the behavior of firms with whom its client may be joined.

Joint and several liability is one focus of current tort reform debates. The limitation of a defendant's liability to its share of the damages also reduces the likelihood of a plaintiff's receiving full compensation, if other defendants are dissolved or bankrupt. Alternative approaches also are being pursued. Since joint and several liability can be viewed as mutual insurance de facto, the chemical industry appears amenable to more formal arrangements.

First, the development of chemical industry standards can reduce the insurer's risk of future damages from the joinder of a careful client to a careless competitor. Because voluntary standards are unenforceable against free riders, the chemical industry generally favors the establishment of tight statutory standards of care. Chemical firms currently inspect common disposal sites to make sure that RCRA standards are being enforced.[11] Insurers also have an incentive to monitor the insured's adherence to regulations, because coverage lapses in the event of willful violation.

Second, in collaboration with environmental organizations, the chemical industry has established a foundation aimed at restoring abandoned multiple-user hazardous waste sites. As suggested by the slow avalanche model, this voluntary cleanup reduces future perils from past practices.[12]

[11]Congressional testimony is summarized in *Chemecology*, the organ of the Chemical Manufacturers Association (CMA); see, for example, "Limit Landfill Use, CMA Spokesman Urges" (February 1983, p. 3) and "Disposal Law May Need Changes, CMA Spokesman Tells Congress" (May 1983, p. 12).

[12]For a description of individual initiatives, see the following in *Chemecology:* "3M Funds Disposal Site Cleanup" (September 1983, p. 10); "Chemical Company [Chevron] Voluntary Action Speeds Disposal Site Cleanup" (November 1983, p. 5); "Monsanto Earmarks $25 Million for 1984 Waste Cleanup Program" (April 1984, p. 8); "Industry Leader [Ciba-Geigy] Urges Voluntary Waste Cleanup" (April 1984, p. 8). For a discussion of the role of large companies in the formation of Clean Sites, Inc., also see *Chemecology:* "Speeding Hazardous Waste Site Cleanup—Industry, Conservationists Work Together" (May 1984, pp. 2–3); "Environmental, Industry Groups Tackle Hazardous Waste Disposal Sites" (July/August 1984, p. 2).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Third, the chemical industry can establish a mutual insurance company or a "captive." Adherence to industry standards of care can become a precondition of insurability. Because the chemical industry is in a better position to calculate its own risks than the insurers, industry-owned mutual insurance pools have a comparative advantage over traditional insurers. The superior technical knowledge of mutuals diminishes the adverse selection problem. Mutuals could contract with conventional insurers to perform claims-settlement and other administrative functions.

The rudiments of a mutual market are visible. An asbestos removal firm has formed a captive that will cover others in the industry (Tarnoff 1985). Hazardous waste management firms have formed a mutual, Waste Insurance Liability, Ltd. Sixteen chemical companies have established Primex, Ltd. to provide excess CGL coverage.[13]

*Alternatives to a Pollution Liability Market*

If the pollution liability market fails to revive spontaneously, there may be considerable political pressure to create artificial markets. For example, states might create assigned risk pools. Insurers that wished to write CGL policies in a state may be required to write EIL policies. Because a particular underwriter might refuse to do business in such a state, assigned risk pools would have to be established in all of the major underwriting jurisdictions simultaneously. Alternatively, the federal government might establish its own insurance program.

Assigned risk pools or federal insurance funds would function in a regulatory capacity only if they were able to set premiums freely. If premiums were regulated on grounds of equity or affordability, then insurance would serve no deterrent function. State automobile-liability pools provide poor examples of flexible premium setting, for most mandate cross-subsidies from good drivers to bad. Similarly, federal deposit, pension, crop, and flood insurance programs have suffered considerable pressures to equalize and subsidize premiums (Mehr and Cammack 1980, p. 257).

A federal pollution insurance program would be subject to intense lobbying on the part of business to obtain similar premium subsidies. Since the principle that "the polluter pays" is firmly entrenched in

[13]"House Panel, Association Announce Efforts to Resolve Environmental Insurance 'Circus'," *Environmental Reporter* 16 (17 January 1986): 1766–67; *Insurance Review* 48 (September 1986): 16.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

public policy, public anxiety over toxic chemicals may eliminate pressures to make premiums "affordable."

If none of these insurance strategies work, the experiment in market-mediated risk management will fail. The remaining approach is to separate the problems of deterrence and victim compensation. Statutory standards of care would continue to serve the deterrent function, badly at that. Either first-party medical and disability insurance or a quasi-public fund, such as workers' compensation, would serve the compensation function.

First-party insurance for medical expenses and property damage is a workable risk-spreading device (Danzon 1984). Insurers would not have to distinguish between environmental, occupational, or other causes. Such injuries as pain and suffering, emotional distress, and birth defects are not insurable on a first-party basis. Requiring victims to pay for their own insurance against pollution-engendered damages, however, would be perceived as unfair by the public.

In contrast, the system of workers' compensation is funded by the employer and accepts "rebuttable presumptions" of causality that lessen the burden of proof and expedite compensation on a no-fault basis. The establishment of workers' compensation was supported by both management and labor as a means of reducing the transactions costs of settlement. As a result, the expected value of the recovery could increase without radically increasing the payments by employers (Friedman and Ladinsky 1967). Since the premiums are based upon the employer's particular loss experience, the system provides positive incentives for careful risk management.

A congressionally mandated study suggested the creation of a two-tiered compensation mechanism for toxic injuries. Tier one consists of an administrative system for reimbursing medical payments and lost wages, based upon rebuttable presumptions. The compensation fund is to be financed by a tax on the production of chemicals, such as that earmarked for Superfund. Tier two consists of the new toxic tort law, with more formidable barriers to recovery balanced against the potential for large awards for pain, suffering, and other damages. Because of the lesser burden of proof, tier one would offer recovery to a larger number of victims than tier two, but at a considerably lower level of compensation (Soble 1977; U.S. Congress 1980).

When Superfund was passed, victim compensation schemes like those outlined in tier one were defeated. Opponents saw the attenuated burden of proof as offering an open-ended entitlement. Citing the experience of the Black Lung Fund for miners (Strader and Sheehe 1981), opponents saw no grounds for excluding anyone with



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the remotest claim of environmental exposure to chemicals (Schmalz 1982).

## Conclusion

The problem of environmental risk management is difficult, and there is no panacea. The most fruitful approach to risk management is developing statutory regulation, tort law, and insurance as a mutually reinforcing tripod. This is not eclecticism for its own sake, for each regime plays an essential role. Statutory financial responsibility requirements increase the chance that the successful plaintiff will not face a bankrupt defendant. Statutory standards of care, supported by the chemical industry, reduce the moral hazard from joint and several liability. A tort system provides a basis for compensation unattainable through regulatory statutes. Under a stable statutory and tort regime, pollution liabilities become more predictable and hence more insurable.

Undoubtedly, environmental risks are far more complex and more difficult to understand than most insurable exposures, and the knowledge base for effective risk management is weak. There are positive signs, however, that the art of environmental risk analysis is developing rapidly under the impetus of market incentives. Based upon solid risk analysis, insurance premiums can serve as powerful incentives for achieving cost-effective risk-reduction decisions.

Would efficient risk-management decisions, arrived at by market signals, be socially acceptable? The lack of public acceptance of some technological hazards, such as toxic chemicals or nuclear power, often dumbfounds engineers and economists, who believe they have "proven" that these technologies are safer than household hazards. In despair, these experts attribute this discrepancy to irrationality, ignorance, or irresponsible interest-group politics.

While the public tends to overestimate the frequency of rare accidents, they have fairly definite perceptions of the attributes of hazardous technologies. The public is particularly averse to technologies that burden the few with substantial costs for the benefit of the many. As a consequence, potentially efficient chemical risk management decisions may be unacceptable unless compensation is actually paid, swiftly and surely to redress the asymmetry.

The new toxic tort regime and financial responsibility requirements make a substantial contribution to victim compensation and thus to reducing public aversion to chemical hazards. The critical element is the creation of a viable competitive market in pollution liability insurance. While no panacea, pollution liability insurance



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

can serve as the keystone of an efficient and equitable environmental risk management system.

## References


Abraham, Kenneth S. *Distributing Risk: Insurance, Legal Theory, and Public Policy*. New Haven: Yale University Press, 1986.

Aickin, R. Malcolm. "The Role of Insurance in Risk Spreading and Risk Reduction." Paper presented at the Conference on Transportation, Storage, and Disposal of Hazardous Materials, International Institute of Applied Systems Analysis, Laxenburg, Austria, July 1985.

Akerlof, George. "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84 (August 1970): 488–500.

All-Industry Advisory Council (AIRAC). "Pollution Liability: The Evolution of a Difficult Market." Oak Brook, Ill., October 1985.

Ames, Bruce N. "Dietary Carcinogens and Anticarcinogens." *Science* 221 (21 September 1983): 1256–64.

Anderson, Eugene R. "The Proposed ISO General Liability Insurance Policy Revision." Paper presented at the annual meeting of the Risk and Insurance Management Society, New Orleans, 18 April 1985.

Best, Robert K., and Collins, James I. "Legal Issues in Pollution-Engendered Torts." *Cato Journal* 2 (Spring 1982): 101–36.

Calabresi, Guido. *The Cost of Accidents*. New Haven: Yale University Press, 1970.

Cheek, Leslie, III. "Risk-Spreaders or Risk-Eliminators? An Insurer's Perspective on the Liability and Financial Responsibility Provisions of RCRA and CERCLA." *Virginia Journal of Natural Resources Law* 2 (Fall 1982): 149–78.

Clean Sites, Inc. "Clean Sites and Private Action: A Plan to Accelerate Private Hazardous Waste Cleanup." Washington, D.C.: Conservation Foundation, May 1984.

Danzon, Patricia M. "Tort Reform and the Role of Government in Private Insurance Markets." *Journal of Legal Studies* 13 (August 1984): 517–49.

Denenberg, Herbert S. et al. *Risk and Insurance*. Englewood Cliffs, N.J.: Prentice-Hall, 1974.

Ehrlich, Isaac, and Becker, Gary. "Market Insurance, Self-Insurance, and Self-Protection." *Journal of Political Economy* 80 (July/August 1972): 623–48.

Fischhoff, Baruch; Slovic, Paul; and Lichtenstein, Sarah. "Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation." *Journal of Experimental Psychology: Human Perception and Performance* 4 (May 1978): 330–44.

Friedman, Lawrence M., and Ladinsky, Jack. "Social Change and the Law of Industrial Accidents." *Columbia Law Review* 67 (January 1967): 50–82.

Harris, Glenn R. et al. "Groundwater Pollution from Industrial Waste Disposal." *Journal of Environmental Health* 44 (May 1982): 287–95.

Hershey, John C., and Schoemaker, Paul J. H. "Risk Taking and Problem Context in the Domain of Losses: An Expected-Utility Analysis." *Journal of Risk and Insurance* 47 (March 1980): 111–32.




LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hogarth, Robin M., and Kunreuther, Howard. "Ambiguity and Insurance Decisions." *American Economic Review* 75 (May 1985): 386–90.

Huebner, S. S.; Black, Kenneth, Jr.; and Cline, Robert S. *Property and Liability Insurance*. Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Kakalik, James S. et al. "Variation in Asbestos Litigation Compensation and Expenses." Santa Monica, RAND Institute for Civil Justice, R-3132-ICJ, 1984.

Katzen, David B. "A Review of Diagnostics and Treatment of Patients Exposed to Low-Level Contaminants." Paper presented at the American Association for the Advancement of Science, Los Angeles, May 1985.

Katzman, Martin T. *Chemical Catastrophes: Regulating Environmental Risk through Pollution Liability Insurance*. Homewood, Ill.: Richard D. Irwin, 1985.

Katzman, Martin T. "Creating Markets for Catastrophic Insurance: Experimental Evidence in the Pollution Liability Line." Resources for the Future, November 1986.

Landes, William M., and Posner, Richard A. "Tort Law as a Regulatory Regime for Catastrophic Personal Injuries." *Journal of Legal Studies* 13 (August 1984): 417–34.

McKean, Roland N. "Enforcement Costs in Environmental and Safety Regulation." *Policy Analysis* 6 (Summer 1980): 265–80.

Mayers, David, and Smith, Clifford W., Jr. "On the Corporate Demand for Insurance." *Journal of Business* 55 (April 1982): 281–96.

Mehr, Robert I., and Cammack, Emerson. *Principles of Insurance*. Homewood, Ill.: Richard D. Irwin, 1980.

Milliman, Jerome. "Can Water Pollution Policy be Efficient?" *Cato Journal* 2 (Spring 1982): 165–96.

Peto, Richard. "Why Cancer?" In *The Resourceful Earth*, chap. 19. Edited by Julian L. Simon and Herman Kahn. New York: Basil Blackwell, 1984.

Priest, George L. "The Invention of Enterprise Liability: A Critical History of the Intellectual Foundation of Modern Tort Law." *Journal of Legal Studies* 14 (December 1985): 461–528.

Rich, Bradford W. "Environmental Litigation and the Insurance Dilemma." *Risk Management* 32 (December 1985): 34–43.

Schmalz, Richard A. "Superfund and Tort Law Reforms—Are They Insurable?" *The Business Lawyer* 38 (November 1982): 175–92.

Schoemaker, Paul J. H., and Kunreuther, Howard. "An Experimental Study of Insurance Decisions." *Journal of Risk and Insurance* 46 (December 1979): 603–18.

Shavell, Steven. "Liability for Harm versus Regulation of Safety." *Journal of Legal Studies* 13 (June 1984): 357–74.

Slovic, Paul et al. "Preference for Insuring Against Probable Small Losses: Insurance Implications." *Journal of Risk and Insurance* 44 (June 1977): 237–58.

Slovic, Paul; Fischhoff, Baruch; and Lichtenstein, Sarah. "Facts and Fears: Understanding Perceived Risk." In *Societal Risk Assessment: How Safe is Safe Enough?*, pp. 181–212. Edited by Richard C. Schwing and Walter A. Albers, Jr. New York: Plenum, 1980.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Smith, Michael L., and Witt, Robert C. "An Economic Analysis of Retroactive Liability Insurance." *Journal of Risk and Insurance* 52 (September 1985): 379–401.

Soble, Stephen M. "A Proposal for the Administrative Compensation of Victims of Toxic Substance Pollution: A Model Act." *Harvard Journal on Legislation* 14 (June 1977): 683–824.

Strader, James D., and Sheehe, Phillip J. "Federal Black Lung: Ten Years of Legislation and Litigation." *The Forum* (American Bar Association) 16 (Winter 1981): 525–55.

Tarnoff, Stephen. "Asbestos Removal Firm to Form Insurer." *Business Insurance* (3 June 1985).

Tyler, Robert M., and Wilcox, Todd J. "Pollution Exclusion Clauses: Problems in Interpretation and Applications under the Comprehensive General Liability Policy." *Idaho Law Review* 17 (Summer 1981): 497–521.

U.S. Congress. "Injuries and Damages from Hazardous Wastes—Analysis and Improvement of Legal Remedies." Report by the Superfund Study Group, 1 July 1982.

U.S. Environmental Protection Agency (EPA). "Pre-Screening for Environmental Hazards—A System for Selecting and Prioritizing Chemicals." EPA-560/1-77-002. April 1977.

U.S. Environmental Protection Agency (EPA). "Federal Financial Requirements for Owners and Operators: A Summary." SW-962, 1982.

U.S. Treasury. "Hazardous Substance Liability Insurance." 1982.

Viscusi, W. Kip. *Employment Hazards: An Investigation of Market Performance.* Cambridge: Harvard University Press, 1979.

Wood, William. "Nuclear Liability after Three Mile Island." *Journal of Risk and Insurance* 48 (March 1981): 450–64.

Zenter, Rene. "Hazards in the Chemical Industry." *Chemical and Engineering News* 57 (5 November 1979): 25–34.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Interest Groups and the Antitrust Paradox

*Bruce L. Benson, M. L. Greenhut, and Randall G. Holcombe*

Economists have generally assumed that the intention of the antitrust laws is to increase economic efficiency. Many observers, however, have noted that the antitrust laws are applied inconsistently and often do not use economic analysis to promote economic efficiency. Judge Robert Bork (1979) referred to this failure of the antitrust laws to promote economic efficiency as the "antitrust paradox," and Peter Asch (1970) called it the "antitrust dilemma." The special interest theory of regulation developed by Stigler (1971) and others assists in understanding the antitrust paradox, because pursuant to it one must not expect antitrust to be applied to benefit the general public.[1]

The special interest view of economic regulation has found its way into evaluations of the antitrust laws.[2] For example, Judge Richard Posner (1969, p. 87) claimed that Federal Trade Commission (FTC) investigations are seldom in the public interest and are undertaken "at the behest of corporations, trade associations, and trade unions whose motivation is at best to shift the costs of their private litigation

---

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

Bruce L. Benson is Associate Professor of Economics at Florida State University, M. L. Greenhut is Alumni Distinguished Professor of Economics at Texas A&M University, and Randall G. Holcombe is Professor of Economics at Auburn University.

[1]Stigler (1985) does not see the special interest theory as completely convincing in the case of antitrust. Some extensions of the special interest theory of government are found in Posner (1974), Peltzman (1976), McCormick and Tollison (1981), Becker (1983), and Holcombe (1985). Some of the many empirical examinations of the theory include Abrams and Settle (1978), Jarrell (1978), Kau and Rubin (1978), McCormick and Tollison (1981), Smith (1982), and Ross (1984). A parallel development to the literature on interest group regulation is the rapidly growing literature on rent seeking. See Tollison (1982) and Benson (1984) for a discussion of the relationship between the two developments and reviews of the relevant literature.

[2]See, for example, Posner (1969), Faith et al. (1982), Weingast and Moran (1983), Benson (1983b), Benson and Greenhut (1986), and High (1984–85).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

to the taxpayer and at worst to harass competitors." The special interest theory of regulation goes a long way toward explaining the antitrust paradox, but there are important differences between antitrust and most other regulatory constraints on business. Moreover, there are many facets of antitrust that are not readily explained by that theory.

Several aspects of antitrust law differentiate it from the types of regulation that are normally considered within the special interest framework. Regulation is normally concerned with one or a few industries, but the antitrust laws are considerably broader in scope. In addition, the courts play a much more visible role in antitrust than they do in regulation. One can legitimately question how laws as broad in scope as the antitrust laws can be the product of special interests. One can further question how special interests could hope to benefit from court decisions that normally are considered to be outside the influence of narrow special interests.[3] These basic issues are evaluated throughout the paper.

Our thesis is that the antitrust laws are a result of a special interest struggle between small and large economic entities seeking changes in the general economic environment rather than the specific favors usually associated with special interest legislation. We explain why a rather general approach was sought for antitrust laws, not the typical goal of many narrowly focused special interests. We also explain Judge Bork's antitrust paradox, both in origin and application, and account for enforcement by both courts and governmental bureaus of the property rights assigned under the antitrust laws.

## The Antitrust Process

There are three main participants in the antitrust process: special interests; legislators; and the commissioners, bureaucrats, and judges who interpret and enforce the statutes. These groups are not mutually exclusive. In particular, bureaucrats and commissioners may also act as interest groups and possess considerable political power that often is used to further their own interests. Stigler (1971) observed that the object of special interest groups is the transfer of wealth, but this is rarely accomplished by a simple transfer of funds from the public treasury, especially in the antitrust arena. Rather, as Benson (1984) noted, wealth transfers are provided through governmental assignments of or transfers of property rights.

[3]Posner (1972), Rubin (1977), Priest (1977), and Holcombe (1983, chap. 9) all suggest a type of invisible hand mechanism leading the courts to reach efficient decisions.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The process of assigning or transferring property rights can be divided into two distinct stages: first, the assignment of property rights; second, the enforcement of the resulting property rights assignment. The first stage is a legislative function; for the assignment to be worth anything, an enforcement mechanism must also be established. These two stages can be thought of as a single object demanded by interest groups because different enforcement levels are possible given any assignment of rights. Thus, one goal of interest groups is to obtain and defend favorable property rights.

When examining antitrust issues there is particular reason for distinguishing between the assignment and the enforcement of rights, because legislators assign antitrust rights but do not enforce them, even though in theory they could. Legislators delegate the enforcement function to agencies (commissions) and to the courts, and the legislature's delegates wield considerable power and independence.[4]

## The Basic Model

Legislators favor special interest groups but never to the extent that the favored group would prefer. They favor the group up to the point where the marginal political benefit received in exchange for the favor equals the marginal cost in terms of lost support from other groups (Peltzman 1976, p. 217). Of course, the favored group would most prefer to receive greater benefits, that is, up to the point where the marginal benefit is zero. Stigler (1971, p. 4), Posner (1974, p. 350), Peltzman (1976), and Becker (1983) concluded that the political exchange in the United States results in an efficient transfer of wealth from a political perspective. There should be no excessive waste or unnecessary inefficiencies caused by such exchange. As Tullock (1965), Eckert (1973), Hilton (1972), Benson (1983b), and Benson and Greenhut (1986) observed, however, the bureaucrats and commissioners to whom the antitrust enforcement power is delegated do not have incentives to behave efficiently. Why the delegation? The answer is because the legislature has a time constraint, and time spent enforcing the rights it has granted reduces the new rights assignments that the legislature can make (Benson 1983b). In addition, the legislature requires the agreement of a large number of people for it to take

[4]In many cases, enforcement agencies have the ability to assign rights, bypassing the legislature. This opens the possibility that an agency might favor an interest group the legislature does not wish to support and tend to corrupt public officials (see Benson and Baden 1985). One would expect that any deviation between legislatively supported interest groups and agency-supported interest groups to be a short-run phenomenon, however, because ultimately the legislature can control the agency through its budget.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

action, so it will be an inefficient organization for accomplishing the enforcement of rights (Erlich and Posner 1974). For these reasons, the legislature delegates the enforcement of the rights it assigns.

The favored interest groups will also want the enforcement function delegated. The theory of bureaucracy suggests that a bureau will produce more output than would be most preferred by the legislature (Tullock 1965; Niskanen 1968, 1971, 1975); and when enforcement of a rights assignment is the output, the result will be over-enforcement (Benson and Greenhut 1986). Because the favored interest will prefer a rights assignment larger than the legislature will grant, over-enforcement by the enforcing agency will benefit the special interest. This explanation applies not only to all types of regulation but to antitrust as well. Special interests will prefer that the enforcement of rights be delegated because the rights will be over-enforced, increasing the special interest's benefits. As Niskanen (1975) observed, agency managers' incentives are closely linked to the size of the agency's budget; and although his model was applied generally to bureaucracy, it fits antitrust as well.[5]

The antitrust model examined in this paper is a straightforward use of Niskanen's bureaucracy model applied to the regulatory environment characterized by Peltzman and others. In general, special interests approach the legislature with demands for favorable property rights assignments. And rights are assigned in a way that establishes a majority for the legislator. After assignment, enforcement is delegated to an agency that over-enforces. If one then considers the agency to be an interest group, one can even support the Stigler-Peltzman predictions of political efficiency. The initial assignment was made due to a political exchange with an interest group; but once the enforcement mechanism is in place, the enforcer becomes another interest group, thus increasing the demand for enforcement. The supposed conflict between the Stigler-Peltzman interest group approach and the Niskanen bureaucracy approach raised by Weingast and Moran (1983) may not exist at all if bureaucrats play an interest group role.

Several vital aspects of the antitrust laws warrant specification at this point. First, the laws are economy-wide rather than focused on just one or a few industries. Second, the laws are couched in very vague language, so that what constitutes a violation is largely a matter of interpretation. Third, the laws involve a commission. Fourth, they involve the courts. The importance of these features distinguishes

[5]Niskanen's early model (1968, 1971) has been subjected to a number of criticisms, several of which were dealt with by Niskanen (1975). See also Benson (1981, 1983a).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

our antitrust model from the bureaucracy model and from Stigler's model of special interest legislation.

In essence, the antitrust laws represent a transfer of rights from large economic entities to small ones, the response of the legislator having been tuned to the times. This is why the laws are economy-wide in scope, rather than being centered on a few industries.

The advantage of having a vague statute is that it could be enforced, not only against those viewed as a present threat to small entities but also against any future threats. In addition, the special interest vagueness allows paradigms to be applied, such as the incipiency doctrine, in which substance can completely disregard the realities of existing competition and economic theory (Greenhut and Ohta 1979; Greenhut et al. 1985). It also explains away (and contradicts) the idea that the antitrust statutes were designed to improve the country's markets and to yield more efficient economic relationships.

What about antitrust enforcement being effected by commissions and courts? Are not appointed commissions less aggressive than the classical bureaucracy model would suggest? Hilton (1972), Eckert (1973), and others have proposed different outputs by commission regulators than by civil service regulators. Essentially, Hilton considered commissioners to be maximizing support, in effect seeking future employment.[6] But maximizing support is certainly consistent with favoring powerful political interest groups. Eckert viewed their salaries as being fixed and not tied to budget size; thus, he characterized commissioners as effort minimizers. It may appear to follow, then, that commissions have no real impact, as Stigler and Friedland (1965) suggested with respect to electric utilities. But commissions enforcing rights that are in the best interest of the regulated industry, as Stigler (1971) suggested is likely, could easily appear to be doing nothing if by doing "something" observers mean "restricting the industry." Furthermore, non-salary perquisites of the office are tied to the size of a commission's operator and level of enforcement. Thus, while the incentives for excessive enforcement may be relatively weak for a commissioner, they still exist, and the general conclusions of Niskanen's view of enforcers hold.

The courts also offer a basic difference vis-à-vis civil servants in the bureaucracy model. One difference is they are less subject to capture than are bureaus. A second difference is that they will be the final authority, because vague legislation fails to identify specific actions that violate the assigned rights. It is our objective to apply

[6]See also Mitnick and Weiss (1974), Russell and Shelton (1974), and Joskow (1974) for arguments similar to Hilton's.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the bureaucracy model to a set of laws, the antitrust laws, that are economy-wide in scope, vague, and enforced by commissioners and the courts.

## Completing the Theory: Ambiguity and the Courts

The Sherman Act was basically inspired by and lobbied for by interest groups made up of small economic entities, particularly from the agricultural sector. Rather than a specific act aimed against, say, machinery companies, the general law could also protect small farmers from large banks or other present or future entities that might threaten them.[7] This same principle applies today with regard to civil rights legislation. Rather than specific laws, a general law to protect civil rights can be enforced by the courts against any present or future violators. The point is that if the special interest group has a large enough constituency and a distinct identity (for example, small economic entity, minority race), a general regulation enforced by the courts can provide more benefits to the interest group than a specific regulation.

While it is true that a series of specific regulations could accomplish the same thing, this would require the special interest to continue returning to the legislature for additional legislation whenever it appeared that a transfer of rights would be beneficial. A general transfer of rights from large economic entities to small ones will continue to benefit the small entities in specific instances that the small entities could not foresee. Because the characteristics of the recipient of the transfer are relatively well defined, this type of transfer will provide more benefits than would a specific regulation. With a general transfer of this type, however, it is not always clear when the law has been violated, so the courts enter as an interpreter of the intent that underlies the law. As long as the courts interpret the intent as being the protection of small economic entities from large ones, the small entities that compose the special interest will benefit from the transfer.

The ambiguity of a general transfer of this type also benefits the recipient, because one can never be certain when the line has been crossed that constitutes a violation. In the case of antitrust, this could cause large economic entities to wield their power carefully, lest they be accused of a violation. Again, this is to the benefit of the small economic units that made up the special interest.

[7]Note, however, that the Interstate Commerce Commission (ICC) act was passed at about the same time to regulate rail rates, in part because of pressure from farmers (and, of course, from the railroads themselves).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Another factor is that the courts are likely to be less influenced by special interests than by a governmental agency. In many circumstances a special interest would prefer enforcement by an agency as a means of increasing the effective transfer of rights, but this may not always be so. When rights are transferred from large economic entities to small ones, there is always the threat that once an enforcing agency is established, the large entities will use their economic power to influence the agency, resulting in a transfer of rights back to them.[8] When rights are being transferred from large entities to small ones, there is good reason for the special interest to prefer the enforcement of rights to take place through the courts. A court enforcer may prevent the large economic entities from capturing the enforcing agency.

In summary, antitrust laws tend to transfer rights from large entities to small ones. There are two reasons for special interests to prefer court enforcement. First, the general transfer can apply to cases not specifically mentioned in the law; court interpretation is required to identify these particular cases. Second, court enforcement will prevent the large economic entities from whom rights were transferred from capturing the enforcing agency to reverse the transfer.

## Interest Groups and Antitrust

The demand for antitrust legislation began building during the 1870s and 1880s (Areeda 1974, p. 44). It involved the formation of many organizations "with revealing names like the National Anti-Monopoly Cheap Freight Railway League" (Neale 1970, p. 12). The primary source of pressure was from farm groups that faced what they perceived to be excessively high rail rates as well as high prices of farm equipment and other manufactured goods. They believed that the high prices were caused by monopoly power exercised in the market and by import tariffs. There was also a general belief among farmers that eastern financiers controlled the credit market and charged them unfairly high interest rates. "Dissatisfaction with manufacturers of farm machinery and other goods, railroads, and eastern financiers became the cry against monopoly" (Areeda 1974, p. 43), and because "the farmers were better endowed with political influence than economic strength, . . . organizations like the National Grange and the National Farmers Alliance insistently demanded

[8]This can be observed today in the Department of Energy. Originally established to control large oil companies, much of its early effects amounted to a transfer from large oil companies to small ones and to the general public. The department's recent programs, however, primarily benefit large energy producers.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

some control of the railways and of monopolies in general" (Neale 1970, p. 12). Neale concluded that "the paramount aim of politicians" in passing the Sherman Act was to meet this demand for action by such organizations.[9]

The business interests that farmers opposed were not without political power themselves, and Congress did not choose full-scale regulation or nationalization of these enterprises due to the political pressure from business (Areeda 1974, p. 44). Thus, the Sherman Act represents congressional attempts to balance the interests of various interest groups, as described by Peltzman (1976). Interest group pressure rather than a desire for economic efficiency led to the passage of the Sherman Act, and Neale (1970, p. 473) noted that the courts have consistently refused to consider economic efficiency issues in judging whether an antitrust violation has taken place, although many Department of Justice and Federal Trade Commission economists stress such issues in their analyses.

The passage of the Sherman Act fits well into the model of special interest legislation. It was clearly understood at the time that farmers demanded some antitrust action and that the passage of the Sherman Act was a response to that demand (see Gordon 1963). It was never intended to produce efficiency; even the name gives away the fact that the law was intended as a transfer of property rights away from trusts. What else could antitrust mean?

The form of the Sherman Act can also be understood within the context of the model. The act was passed to benefit the farmers as a special interest group, but the businesses the act opposed also had considerable political power. In balancing the interests of both groups, Congress could hardly legislate away the businesses. Viewed in this light, the vague nature of the statute and the delegation of its enforcement to the Department of Justice and to the courts make good sense from the standpoint of the farmers for whom the act was passed. Because the act generally declares combinations, contracts, or conspiracies in restraint of trade to be illegal, it is simply a vague piece of legislation that endorsed the common law and was designed to protect small economic entities from being harmed by large ones. Because the farmers were small entities and the railroads, banks, and manufacturers were large entities, the act appears to be a one-way transfer of property rights to the farmers (and any other small businesses that might feel harmed).

[9]Baxter (1979) and Katzman (1980) have argued that there is no strong evidence that one group is being favored by the antitrust laws; but Williamson (1979), commenting on Baxter's argument, found room for disagreement.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The structure of the act benefited the farmers in several ways. First, because the act did not name offenders ahead of time (it was not aimed at railroads or banks specifically but at anyone acting to restrain trade), action could be taken not only against present offenders but also against future ones. Because individual farmers could always foresee being economically small entities in relation to the other businesses with whom they dealt, the umbrella of the act protected farmers from a wide variety of economic threats.

Second, because of the act's vagueness, it enabled small economic entities to bring complaints under the act in a wide variety of circumstances. That vagueness would also ensure that there was always the possibility of a violation, having the corollary effect of making larger businesses behave cautiously even before any enforcement proceedings began, simply because of the threat of action.

Third, because victims could bring complaints, the Sherman Act could be over-enforced, even though an agency does not have the direct power of enforcement.[10] As a result, the act ended up providing more benefits to the special interests that the act favored than Congress had originally intended. The act was an example of special interest legislation and was recognized as such at the time of its passage.

The next major development in antitrust was the passage of the Federal Trade Commission Act and the Clayton Act in 1914. The establishment of the FTC was the result of many sources of political pressure, including the big businesses that found themselves constrained by the vagueness of the Sherman Act. There was considerable pressure from business organizations for more clearly articulated standards, and some observers have concluded that the establishment of the FTC was a victory for those businesses that wanted to ensure their political power and protect themselves from competition (Areeda 1974, pp. 47–48; Kolko 1963). In this view, the FTC was created to advise businessmen, to approve of their collusive organizations, and to create order in markets. As in the model above, large businesses could not effectively capture the courts, but the FTC gave them an agency through which they could use their economic power and influence.

Clearly, the FTC is not solely a pro-business agency, but was developed through compromise. The *Ralph Nader Study Group*

[10]Neale (1970, pp. 374, 385) noted that both the Federal Trade Commission and the Department of Justice rely almost exclusively on complaints to determine when the antitrust laws should be enforced, and that complaints predictably come from those who believe they are being injured.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*Report on Antitrust* recognized that "in 1914 both sides—those advocating a kind of business advisor and those seeking more energetic trustbusting—compromised to produce the Federal Trade Commission" (Green et al. 1972). And Areeda (1974, p. 48) noted that "similar differences of opinion were reflected in the Clayton Act, passed in the same year. . . . These differences were compromised in the ultimate enactment." As Peltzman's theory of regulation would predict, the favored group will not receive the maximum possible benefit. The FTC and Clayton acts, however, can clearly be seen as a transfer of rights to big business in response to the larger-than-intended transfer away from big business that had been the result of the Sherman Act.

It is interesting that the 1914 rights transfer was reversed somewhat as the FTC and the courts began to rule against incipient violations. In *Triangle Conduit and Cable Company v. FTC* (168 F.2d 175, 7th Cir. [1948]), for example, the court held that manufacturers utilizing a basing-point pricing system in order to stabilize competition among themselves were in violation of the Sherman Act. In itself, this decision seems unremarkable.[11] But the court proposed that unilateral adoption of a basing-point pricing scheme by an individual firm could also be prohibited, as long as other firms might adopt similar schemes. Even if unilateral at the start, the court saw this as the first step toward a conspiracy (see Greenhut [1970] 1974, chaps. 7, 14). In *Fashion Originators Guild v. FTC* (381 U.S. 357, 367 [1965]), the Supreme Court held that it was an object "of the Federal Trade Commission Act to reach not merely in their fruition but also in their incipiency," and in *E. B. Muller and Co. v. FTC* (142 F.2d 511, 517 6th Cir. [1944]) the Sixth Circuit Court ruled that "the purpose of the Federal Trade Commission Act is to prevent potential injury by stopping unfair methods of competition in their incipiency." Clearly in antitrust cases this represents a transfer of rights from large economic entities to small ones. The party bringing the complaint has to show neither actual nor present harms; it must only show that some harm can be predicted as a result of the offending policy.[12]

## Exemptions and Interest Groups

Perhaps the strongest evidence of the influence of interest groups on antitrust legislation lies in their success in gaining immunity to

[11]See Greenhut ([1956] 1983) for a concurrence with the decision.

[12]The incipiency doctrine reflects the public's distrust of oligopolies as well as the general failure (including that of economists) to distinguish between organized (collusive) and competitive oligopolies, the market prices and factor incomes of which are



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the antitrust laws (Adams 1965). Agricultural interests comprised the major pressure group pushing for the early antitrust laws, so it is not surprising that agricultural organizations were able to obtain specific exemptions from coverage under the antitrust laws. Section 6 of the Clayton Act partially exempts agricultural organizations, allowing farmers to form cooperative associations without violating the law. The Capper-Volstead Act of 1922 extends the Clayton Act to exempt capital stock agricultural cooperatives that were not exempted under the Clayton Act.

The courts have interpreted the Clayton Act as allowing farm organizations to "set association policy, fix prices at which their cooperative will sell their produce, and otherwise carry on like a business corporation without thereby violating the antitrust laws."[13] Thus, agricultural interest groups—who were instrumental in having the antitrust laws passed as a transfer of property rights to them—also had the political power to see that the laws were not applied to them, even when their actions represented combinations or contracts in a restraint of trade that clearly would have been considered a violation if committed by another business entity. A similar exemption was given to fishermen's organizations in the Fisheries Cooperative Marketing Act of 1934. In some cases, the courts ruled in favor of the special interests; in others the interests were able to get additional acts passed that specifically exempted them.

In 1908, the Supreme Court ruled in *Lowe v. Lawlor* (208 U.S. 274) that a nationwide boycott organized by a union to persuade wholesalers and retailers not to buy a particular firm's product was an interference with the interstate shipment of goods and, therefore, a restraint of trade. The Court awarded treble damages and ordered the union and individual union members to pay shares of the award. Following the decision, labor union officials "immediately commenced pressure for exemption of labor from the antitrust laws." The drive resulted in the Clayton Act's declaration that labor organizations are exempt from the antitrust laws (Northrup and Bloom 1965, p. 313). Section 6 of the act reads: "[N]othing contained in the antitrust laws shall be construed to forbid the existence and operation of labor . . . organizations, instituted for the purposes of mutual help,

---

of opposite order (Greenhut [1970] 1974). This doctrine and its increasing use, reflected even in Judge Learned Hand's *Alcoa* decision, are based on the failure to recognize that conscious parallelism of action by oligopolists can derive from competitive behavior and in turn generate market-efficient results. The vested interests of those who identify with the small entities must condemn any form of Loschian conjectural variations behavior (Greenhut et al. 1975) and thus relegate the antitrust laws to a set of restrictions with no roots at all in any quest for economic efficiency.

[13]Maryland and Virginia Milk Producers v. United States, 362 U.S. 458, 466 (1970).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and not having capital stock or conducted for profits, or to forbid or restrain individual members of such organizations from lawfully carrying out the legitimate objects thereof, be held or construed to be illegal combinations or conspiracies in restraint of trade under the antitrust laws." Section 20 of the act prevents the use of federal injunctions against strikes, boycotts, picketing, and similar activities "in any case between an employer and employees, or between persons employed and persons seeking employment, involving or growing out of, a dispute concerning terms or conditions of employment."

Despite the Clayton Act's exemption of labor from the antitrust laws, the Supreme Court ruled in 1921 in *Duplex Printing Company v. Deering* (254 U.S. 433) that labor unions could be held accountable under the antitrust laws for some of their actions. In this case, a union that was trying to organize the Duplex plant had succeeded in getting members of other unions to refuse to handle the firm's products. The Court held that Section 20 protected a union only when its members were employed by the company they were acting against. The American Federation of Labor turned to Congress; after continual pressure, the Norris-La Guardia Act was passed in 1932, which deprived the federal courts of antitrust jurisdiction in almost all labor disputes.

The labor case parallels the agricultural case. Though labor interests were not actively seeking passage of the antitrust laws, this powerful lobbying group nevertheless managed to remain exempt from the laws' influence. If the courts did not rule in favor of a special interest, then the special interest was able to turn to Congress to have exemptions to the antitrust laws legislated for it.

Other exemptions could be mentioned. For example, if regulated industries are the beneficiaries of regulation, as Stigler (1971) argued, then the regulated firms should be able to protect their benefits from potential challenges through the antitrust laws. As Areeda (1974, pp. 105–14) and Adams (1965, pp. 277–84) noted, regulated industries are largely exempt from the antitrust laws. Similarly, the Miller-Tydings Act (1937) and the McGuire-Keough Act (1952) were results of a movement among small retailers that had successfully obtained passage of "fair trade" laws in 45 states prior to 1937. This legislation exempted firms from antitrust laws in states with legalized minimum resale prices. The "fair trade" laws, which are widely viewed as protecting small shopkeepers from the more efficient larger retail chains (in a clearly collusive manner in restraint of trade), caused Neale (1970, p. 276) to note that "many political groups which would yield to none in zeal for trust busting are to be found in the van of the so-called fair trade movement. . . ." The Robinson-Patman Act preventing price discrimination is yet another example of an antitrust



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

law designed to protect a special interest, that interest once again being small retailers who could not buy in large enough quantities to receive quantity discounts often given to chains.[14] The exemptions to the antitrust laws are all easily recognized as the result of special interest pressure on the legislature.

## The Antitrust Laws over Time

Neale (1970, p. 11) observed that "there is evidence that the aims and scope of antitrust policy have changed a good deal since the passage of the Sherman Act, and may easily change some more in the future." Following the interest group theory of regulation, this is precisely what one would expect if the relative strengths of interest groups change over time (Benson 1984; Weingast and Moran 1983). For example, Clarkson and Muris (1982) noted that prior to 1969 the bulk of the FTC antitrust enforcement was aimed at discouraging price competition under Robinson-Patman when it threatened the well-being of small firms. After significant criticism, the FTC in 1969 suddenly closed about 600 of its investigations, reorganized the commission's work to deemphasize Robinson-Patman, and began several large-scale industry investigations. Similarly, the settling of the AT&T case and the closing of the IBM case within months of each other can be seen as a response to a change in the political climate during the Reagan era.

Posner's (1969, p. 83) study of the FTC emphasizes that a congressman must support the demands of interest groups in his district, noting that "the welfare of his constituents may depend disproportionately on a few key industries. The promotion of the industries becomes one of his most important duties as a representative of the district." One might well expect cases to be biased in favor of the firms in the districts of legislators serving on committees that have oversight over the FTC; this is exactly what Faith et al. (1982) found. Weingast and Moran (1983, p. 775) reached a similar conclusion with regard to oversight committees, noting that "markedly different preferences on the committee lead to major shifts in agency policy." Because oversight committees are merely a reflection of interest group demands, Weingast and Moran and Faith et al. conclude that the FTC's output is best explained by the interest group theory.

[14]See Ross (1984) for a discussion of the interest groups who were influential in obtaining passage of the Robinson-Patman Act and for empirical estimates of the benefits obtained and the costs incurred by various groups as a result of the enforcement of the act.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

This discussion places a different light on the general vagueness of the antitrust laws. Earlier, the vagueness of the laws was seen as a means to transfer rights to one well-defined group (for example, small economic entities) from other unspecified groups. Small farmers could obtain transfers from railroads, banks, manufacturers, and any other future economic threats without the threats being spelled out in the law. But the laws' vagueness also benefits the legislature; due to its oversight capacity, the legislature will be able to influence the type of enforcement in response to the changing power and demands of interest groups, all without writing a new law.

In light of the fairly obvious political nature of the passage and enforcement of antitrust laws, we might wonder why such widespread acceptance of the efficiency-enhancing goals of antitrust appears to characterize the economics profession. This view appears to be an after-the-fact rationalization, however, because economists were not advocates for or even concerned with passage of the earliest statutes. As Scherer (1970, p. 424) reported:

> About the only group in America other than big businessmen outspokenly unconcerned about the trust problem were the professional economists. Many were captivated by Darwin's theory of biological selection. They saw the growth of big business as a natural evolutionary response consistent with economies of scale, or when economics were patently absent from mergers, as a step necessary to eliminate cut-throat competition. But in that unenlightened era, the views of unenlightened economists concerning big business had little influence on public policy.

Clearly, Scherer holds with the efficiency-enhancing perception of antitrust, and he reflects the view that came to dominate the profession early in the 20th century.

## Summary and Conclusion

This paper has fit the antitrust laws into the special interest theory of regulation, stressing the influence of special interests on antitrust. The paper also spotlighted the significant differences between antitrust and other forms of regulation, emphasizing the vagueness of the antitrust statutes, their broad focus, and the fact that the laws are generally subject to interpretation and enforcement by the FTC and the courts rather than by a narrowly focused agency. It is clear that the effects of antitrust diverge from those of other forms of regulation. For example, the interaction of vagueness and use of the courts sometimes requires legislators to pass new statutes in order to protect the favored special interest. On the other hand, to the extent that the



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

FTC serves as the final authority, the legislators' interest in reelection is being protected by insulating the legislature from direct responsibility for specific actions. Most generally, we propose that antitrust in the United States centers on special interest effects and the legislators' quests for reelection. The real antitrust paradox is to have expected legislatures to be concerned with the general welfare.

## References


Abrams, Burton A., and Settle, Russell F. "The Economic Theory of Regulation and Public Financing of Presidential Elections." *Journal of Political Economy* 86 (April 1978): 245–57.

Adams, Walter. "Exemptions from Antitrust: Their Extent and Rationale." In *Perspective on Antitrust Policy*, pp. 273–311. Edited by Almarin Phillips. Princeton, N.J.: Princeton University Press, 1965.

Areeda, Phillip. *Antitrust Analysis*. Boston: Little, Brown and Co., 1974.

Asch, Peter. *Economic Theory and the Antitrust Dilemma*. New York: John Wiley and Sons, 1970.

Baxter, William F. "The Political Economy of Antitrust." In *The Political Economy of Antitrust: Principal Paper by William Baxter*, pp. 3–50. Edited by Robert O. Tollison. Lexington, Mass.: Lexington Books, 1979.

Becker, Gary S. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (August 1983): 371–400.

Benson, Bruce L. "Why Are Congressional Committees Dominated by 'High Demand' Legislators?" *Southern Economic Journal* 48 (July 1981): 68–77.

Benson, Bruce L. "High Demand Legislative Committees and Bureaucratic Output." *Public Finance Quarterly* 11 (July 1983a): 259–81.

Benson, Bruce L. "The Economic Theory of Regulation as an Explanation of Politics Toward Bank Mergers and Holding Company Acquisitions." *Antitrust Bulletin* 28 (Winter 1983b): 839–62.

Benson, Bruce L. "Rent Seeking from a Property Rights Perspective." *Southern Economic Journal* 51 (October 1984): 388–400.

Benson, Bruce L., and Baden, John. "The Political Economy of Government Corruption: The Logic of Underground Government." *Journal of Legal Studies* 14 (June 1985): 391–410.

Benson, Bruce L., and Greenhut, M. L. "Special Interest, Bureaucrats, and Antitrust: An Explanation of the Antitrust Paradox." In *Antitrust and Regulation*, pp. 53–90. Edited by Ronald E. Grieson. Lexington, Mass.: Lexington Books, 1986.

Bork, Robert H. *The Antitrust Paradox: A Policy at War with Itself*. New York: Basic Books, 1978.

Clarkson, Kenneth, and Muris, Timothy J. "Letting Competition Serve Consumers." In *Instead of Regulation: Alternatives to Federal Regulatory Agencies*, pp. 135–68. Edited by Robert W. Poole. Lexington, Mass.: Lexington Books, 1982.

Eckert, Ross. "On the Incentives of Regulators: The Case of Taxicabs." *Public Choice* 14 (Spring 1973): 83–99.




LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Erlich, Isaac, and Posner, Richard. "An Economic Analysis of Legal Rule-Making." *Journal of Legal Studies* 3 (January 1974): 257–86.

Faith, Roger L.; Leavens, Donald R.; and Tollison, Robert D. "Antitrust Pork Barrel." *Journal of Law and Economics* 25 (October 1982): 329–42.

Gordon, Sanford D., "Attitudes Towards Trusts Prior to the Sherman Act." *Southern Economic Journal* 30 (October 1963): 156–67.

Green, Mark J.; Moore, Beverly C.; and Wasserstein, Bruce. *The Closed Enterprise System: Ralph Nader's Study Group Report on Antitrust Enforcement*. New York: Grossman Publishers, 1972.

Greenhut, M. L. *Plant Location in Theory and Practice* [1956]. Westport, Conn.: Greenwood Publishing Co., 1983.

Greenhut, M. L. *A Theory of the Firm in Economic Space* [1970]. Austin, Texas: Lone Star Publishers, 1974.

Greenhut, M. L., and Ohta, H. "Vertical Integration of Successive Oligopolists." *American Economic Review* 69 (March 1979): 137–41.

Greenhut, M. L.; Hwang, C. S.; Norman, G.; and Smithson, C. "An Anomaly in the Service Industry: The Effect of Entry on Fees." *The Economic Journal* 95 (March 1985): 169–77.

Greenhut, M. L.; Hwang, C. S.; and Ohta, H. "Observations on the Shape and Relevance of the Spatial Demand Function." *Econometrica* 63 (July 1975): 669–82.

High, Jack. "Bork's Paradox: Static versus Dynamic Efficiency in Antitrust Analysis." *Contemporary Policy Issues* 3 (Winter 1984–85): 21–34.

Hilton, George. "The Basic Behavior of Regulatory Commissions." *American Economic Review* 62 (May 1972): 47–54.

Holcombe, Randall G. *Public Finance and the Political Process*. Carbondale: Southern Illinois University Press, 1983.

Holcombe, Randall G. *An Economic Analysis of Democracy*. Carbondale: Southern Illinois University Press, 1985.

Jarrell, G. "The Demand for State Regulation of the Electric Utility Industry." *Journal of Law and Economics* 21 (October 1978): 269–95.

Joskow, P. "Inflation and Environmental Concern: Structure Change in the Process of Public Utility Price Regulation." *Journal of Law and Economics* 17 (October 1974): 291–327.

Kau, James B., and Rubin, Paul H. "Voting on Minimum Wages: A Time Series Analysis." *Journal of Political Economy* 82 (April 1978): 337–42.

Katzman, Robert A. *Regulatory Bureaucracy: The Federal Trade Commission and Antitrust Policy*. Cambridge, Mass.: MIT Press, 1980.

Kolko, Gabriel. *The Triumph of Conservatism: A Re-Interpretation of American History, 1900–1916*. New York: Free Press, 1963.

McCormick, Robert E., and Tollison, Robert D. *Politicians, Legislation, and the Economy: An Inquiry into the Interest Group Theory of Government*. Boston: Martinus Nijhoff Publishing, 1981.

Mitnick, B., and Weiss, C. "The Siting Impasse and a Rational Choice Model of Regulatory Behavior: An Agency for Power Plant Siting." *Journal of Environmental Economics and Management* 1 (August 1974): 150–71.

Neale, A. D. *The Antitrust Laws of the U.S.A.* Cambridge: Cambridge University Press, 1970.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Niskanen, William A. "The Peculiar Economics of Bureaucracy." *American Economic Review* 58 (May 1968): 293–305.
Niskanen, William A. *Bureaucracy and Representative Government.* Chicago: Aldine-Atherton, 1971.
Niskanen, William A. "Bureaucrats and Politicians." *Journal of Law and Economics* 18 (December 1975): 617–43.
Northrup, Herbert R., and Bloom, Gordon F. "Labor Unions and Antitrust Laws: Past, Present, and Proposals." In *Perspectives in Antitrust Policy,* pp. 312–54. Edited by Almarin Phillips. Princeton: Princeton University Press, 1965.
Peltzman, Sam. "Toward a More General Theory of Regulation." *Journal of Law and Economics* 19 (August 1976): 211–40.
Posner, Richard. "The Federal Trade Commission." *University of Chicago Law Review* 37 (1969): 47–89.
Posner, Richard. *Economic Analysis of Law.* Boston: Little, Brown, 1972.
Posner, Richard. "Theories of Economic Regulation." *Bell Journal of Economics and Management Science* 5 (Autumn 1974): 335–58.
Priest, George L. "The Common Law Process and the Selection of Efficient Rules." *Journal of Legal Studies* 6 (January 1977): 65–82.
Ross, Thomas W. "Winners and Losers under the Robinson-Patman Act." *Journal of Law and Economics* 27 (October 1984): 243–71.
Rubin, Paul H. "Why Is the Common Law Efficient?" *Journal of Legal Studies* 6 (January 1977): 51–63.
Russell, Milton, and Shelton, Robert. "A Model of Regulatory Agency Behavior." *Public Choice* 20 (Winter 1974): 47–62.
Scherer, F. M. *Industrial Market Structure and Economic Performance.* Chicago: Rand McNally, 1970.
Smith, Janet K. "Production of Licensing Legislation: An Economic Analysis of Interstate Differences." *Journal of Legal Studies* 11 (January 1982): 117–37.
Stigler, George J. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2 (Spring 1971): 3–21.
Stigler, George J. "The Origin of the Sherman Act." *Journal of Legal Studies* 14 (January 1985): 1–12.
Stigler, George J., and Friedland, Claire. "What Can Regulators Regulate? The Case of Electricity." *Journal of Law and Economics* 5 (October 1965): 1–16.
Tollison, Robert D. "Rent Seeking: A Survey." *Kyklos* (1982): 575–602.
Tullock, Gordon. *The Politics of Bureaucracy.* Washington, D.C.: Public Affairs Press, 1965.
Weingast, Barry R., and Moran, Mark J. "Bureaucratic Discretion or Congressional Control? Regulatory Policymaking by the Federal Trade Commission." *Journal of Political Economy* 91 (October 1983): 765–800.
Williamson, Oliver. "Commentary." In *The Political Economy of Antitrust: Principal Paper by William Baxter,* pp. 77–93. Edited by Robert D. Tollison. Lexington, Mass.: Lexington Books, 1979.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Interstate Trade Barriers and the Constitution

*Steven G. Craig and Joel W. Sailors*

The existence of substantial barriers to interstate trade is one of the more surprising aspects of the public policy environment. Barriers to trade are generally familiar to students of public policy from an international perspective, but barriers to international trade are a small part of the trade problem facing the United States. A more serious problem, and one that is entirely within the purview of U.S. policymakers, is the extensive level of trade barriers between states. This paper explores the pattern of interstate trade barriers and illustrates the extent to which they permeate the economy. Given the level of trade barriers that are uncovered, we examine the policy environment that has allowed the Balkanization of national markets to develop. We focus particularly on the Supreme Court, which has failed to protect the national economy from the actions of individual state protectionist policies.

Evidence on the extent to which individual states interfere with trade from other states apparently has eluded the major economic policymakers in the federal government. For example, the 1985 *Economic Report of the President* (p. 115) declares:

> The debilitating effects of . . . protectionism on the States' economies convinced the framers of the U.S. Constitution to forbid individual States from levying tariffs. . . . The constitutional ban on State tariffs was crucial to the development of the U.S. economy not only because it established a free-trade area among the 13 original States, but also because it ensured that the free-trade area would expand automatically as new States joined the Union.

---

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

Steven G. Craig is Assistant Professor of Economics and Joel W. Sailors is Professor of Economics at the University of Houston. They acknowledge the financial support of the Center for Public Policy and the helpful comments on an earlier version of this paper by Janet Kohlhase, David Papell, two anonymous referees, and the editor.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Our evidence indicates that the economy grew despite the heavy burden states have imposed on trade with each other. The total value of interstate trade in manufacturing alone has been estimated to be almost equal to the U.S. GNP, about 18 times the value of U.S. exports (Rodgers 1973).[1] Because of the magnitude of interstate trade, economic distortions from seemingly small interferences in trade can result in large losses.

In order to assess the effectiveness of the current policy environment, we examine potential motivating factors that would explain the current pattern of interstate trade restrictions. Trade barriers occur primarily because of the federalist structure of the United States, under which each state has autonomous power independent of the federal government. This structure is a source of strength for our nation, as it allows each state to pattern policies to best fit its local situation. It is also possible, however, for federalism to cause problems for the economy, as each state will act to serve what it perceives as its best interest. On some occasions, active pursuit of an individual state's interest is inconsistent with the best interests of the country. Even more important, individual states may pursue policies that benefit a minority of its residents at the expense of the majority. This in-state redistributive behavior, which is only permissible under a federalist system, is severely damaging the national economy by distorting the trade patterns that otherwise would occur.

Whatever the stated purpose for erecting barriers to interstate trade, the primary effect is to protect the home-state markets from out-of-state competition. A reduction in the number of competitors allows in-state suppliers to charge consumers higher prices; thus, the firms receive higher profits. The dominant form of trade barriers is administrative restrictions, such as licensing, prohibition on trade, or costly performance criteria. Except in a small number of industries, import taxes (tariffs) on interstate trade have been declared unconstitutional by the Supreme Court. This restriction on the form of trade barriers is not important to economic efficiency, although the form of trade restrictions (administrative restrictions or tariffs) is important for determining what share of the economic cost is borne by consumers. Unfortunately, the primary focus of 200 years of Supreme Court attention has been on the *form* of interstate trade barriers. As yet, there has been no effort to systematically reduce the degree of economic distortion resulting from interstate barriers to trade.

[1]Interstate trade in manufactures for the year 1963 is estimated to be $580.4 billion, compared to a GNP of $589.2 billion (Rodgers 1973). Manufacturing accounted for 61.5 percent of GNP in that year (see infra, n. 2).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Barriers to trade with other states are ostensibly prohibited by the Commerce Clause of the U.S. Constitution. This clause essentially mandates that no state shall take any action that inhibits trade with any other state. While the clause is clear on the surface, its interpretation has allowed a large number of trade barriers to persist. We explore the cause of the failure of the Commerce Clause and present evidence indicating that the Supreme Court has had difficulty disentangling the issues involved.

## Interstate Trade Barriers

Interstate trade is crucial to the healthy operation of the U.S. economy. The central goal of the Commerce Clause is to ensure that the United States is internally a free trade zone, which maximizes the gains from trade to the country as a whole. Despite the existence of substantial barriers to interstate trade, trade between states is an extremely large part of the economy. In an extensive and unique study, Rodgers (1973) determined the size of interstate trade in the manufacturing sector of the economy for the year 1963. To update his estimate to 1980, assume that the same proportion of all manufactured goods crosses state lines in 1980 as in 1963. This yields an estimate of interstate trade in manufactured goods of $2.2 trillion. This estimate is equal to 85 percent of GNP, because much of interstate trade is in intermediate goods in the production process rather than in the final goods that GNP measures. Rodgers's estimate of interstate trade does not include the service sector of the economy; but trade in services, such as in banking, insurance, labor services, and tourism is also important. It is clear that the interstate barriers to trade are central to the overall health of the economy, as they affect trillions of dollars worth of transactions.

There is no current estimate of the amount of trade that has been prevented due to the existing levels of trade barriers, but an additional barrier in the manufacturing sector of even one-tenth of 1 percent would amount to a tax of $2.2 billion on the economy. As a comparison, Magee (1972) determined that the economic loss to the United States from trade barriers it imposed on international trade is 8.8 percent ($197 billion) of the value of imports.[2] Over 80 percent of the economic loss calculated by Magee is the result of administrative restrictions on trade, also known as nontariff barriers to trade.

[2]This $197 billion loss represents only half of the potential loss, since losses to "foreigners" are also suffered in the United States with interstate trade. It should also be noted that the service sector, now over half of GNP, has been excluded from these loss calculations.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Under current U.S. trade policy, these nontariff barriers are the primary form of trade restriction.

There are two primary types of interstate trade barriers: export taxes and import restrictions. Export taxes are levied on goods and services that are consumed mainly by people who reside outside the state that levies the tax. Import restrictions take two principal forms: taxes and administrative barriers. In general, import taxes at the state level are illegal. Because of this limit on the form of import trade barriers, the main types of interstate trade restrictions are administrative barriers, which are an important source of economic distortions.

Interstate trade barriers are difficult to count, as each state has a large number and variety of restrictions on trade. Nevertheless, we have cataloged trade barriers in many major areas of the economy, including agriculture, professional labor, banking, insurance, and state governmental purchases. These categories help to show how extensive trade barriers are and illustrate the various forms they can take.

Agriculture suffers from a complex array of administrative restrictions. Hillman and Rowell (1952) found over 1,500 laws in only 11 western states that impede efficient agricultural trade between states. Among the products affected are dairy products, oleomargarines, livestock, eggs, general foods, and nursery stock. The form of the laws varies considerably between states; it affects grades, standards, labeling, motor transportation, quarantine, and inspection. From 1925 to 1973, for example, California residents could not purchase Florida-grown avocados. The California law mandated that avocados have at least an 8 percent oil content to be considered mature, and Florida avocados typically contain less oil than that required. Similarly, Texas residents could not purchase Florida-grown grapefruit from 1967 to the mid 1970s because the Florida grapefruit was declared inedible due to insufficient sugar. Agricultural trade restrictions have many purposes, some of which are to restrict the spread of harmful disease and insects and to protect consumers from poor quality. But most legislation does not pursue these goals in a cost-effective manner, primarily because the goal of the legislation is to protect in-state growers by restricting competition within their home markets.

Professional labor certification and licensing are a common form of restricting interstate commerce. Labor certification restricts trade through a monopolization effect, which occurs when entry into a profession is made more difficult and wage rates are driven up due to the lack of competition, and through an effect that arises because state licensing requirements vary among states. Strictly interpreted, it is this second effect that is the interstate trade barrier. But the



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

monopolization effect is also a crucial barrier to trade because a potentially important source of competition to professionals in one state are the practitioners from other states who may move if they perceive the possibility of higher income. As a result, there is substantial incentive for in-state professionals to resist reciprocity of certification among states. It is not clear just how much certification would exist if interstate reciprocity was prevalent; even the monopolization effect may be driven by the desire to restrict the interstate movement of professionals.

In her comprehensive study, Greene (1969) found more than 2,800 different state laws affecting labor by a wide variety of professionals, including doctors, lawyers, dentists, other medical professionals, and teachers. The interstate differences in certification laws sharply restrict the workers in these professions from moving to other states, making it difficult for workers to move to states where their services are most highly rewarded. Several studies have found that certification reduces the competition faced by those in high-income states. Shepard (1978) found that dentists have about 12 percent more income than they would if they operated in a market without restrictions, a condition that cost consumers $700 million in 1970. Over half of the economic loss is directly attributable to interstate variation in certification procedures. Similarly, Benham and Benham (1975) found that legal control over optometrists in some states caused eyeglasses to be 25 to 40 percent more expensive than they were in states with fewer restrictions. While the mobility restrictions do not entirely prohibit interstate movement of professionals, they make such movement more costly. If prices are higher in one state than another, it may indicate that there are greater potential profits in the high-price state. To the extent that there are higher potential profits, professionals in other locations will tend to move to the high-price state, eventually eliminating the price disparity through competition. The importance of migration, although by consumers rather than providers, is shown directly by Benham and Benham (1975). They found that residents living near state lines of less-restrictive states paid lower prices than did other residents of the restrictive states. Presumably, this is due to the greater level of competition provided near the state line by the less-restrictive state.

The banking industry has benefited from state protection since 1956, when the Bank Holding Company Act made cross-state banking illegal without express state permission. Until 1982 only one state, Maine, had allowed interstate banking. Since then, there has been some relaxation, although the issue is by no means settled. Several New England states have instituted inter-regional banking laws, but



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

they still prohibit the entrance of banks from outside the region. One fear about interstate banking is that large, multi-state banks will drive smaller, less cost-effective banks out of business. Recent studies have found that large banks do not appear to have much of a cost advantage, but they also found that profits in a more competitive environment are significantly lower (Whitehead and Luyties 1984). That is, prices to consumers are lower in the more competitive environment. These advantages to local banking consumers have been prevented completely until the recent deregulation and are only beginning to be available.

Interstate trade restrictions have hampered the insurance industry for a number of years. Texas maintained a set of complicated restrictions from 1908 that discriminated against out-of-state insurance companies. These regulations initially prohibited entry by out-of-state firms, although they were gradually eroded until they were repealed in 1963. Unlike most industries, states have been allowed to explicitly levy taxes on insurance firms based out of state while exempting in-state firms from the tax (Benson 1970).[3] As a result, several states have differential premium taxes on out-of-state firms. For example, Michigan currently taxes out-of-state firms 2 percent of the premiums they collect and exempts in-state companies. Again, the focus of these restrictions has been to protect local firms at the cost of increased prices for state residents.

Some state governments have granted special privileges to resident firms that provide goods and services to the state government. These privileges take the form of purchasing preference laws that allow in-state firms to bid a certain percentage (usually 5 percent) over an otherwise equal out-of-state firm and still be awarded a state government contract. A special form of the percentage preference law is a reciprocal preference law, which only grants the preference if the out-of-state firm is located in a state that has its own percentage preference law. We examined these purchasing preference laws extensively in an earlier study (Craig and Sailors 1984) and found that they cause a significant increase in state governmental spending.[4] There are 9 states with percentage preference laws and 11 states with reciprocal preference laws. The average state with a percentage

[3]Benson (1970) showed that insurance was declared not a "good" subject to the Commerce Clause because it is a service industry. It is not clear from a layman's perspective why the distinction between goods and services might matter. Congress later overruled the Supreme Court when the Court reversed itself on this issue.

[4]Purchasing preference laws have the same impetus as national trade barriers, namely, employment protection. For example, during the 1930s 26 states enacted public purchasing preference laws (Melder, n.d.). Also see Massel (1962).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

preference law spent $10.13 in real terms per capita more than an otherwise similar state. This amounts to 3 percent of state per capita expenditure in states with these laws and equaled $1.6 billion in 1980 alone. The nine states with the percentage preference laws thus extracted $1.6 billion in 1980 from their taxpayers in order to provide extra payments to firms located in the state. We found that state per capita income would have to rise 32.3 times the extra expenditure to generate enough tax revenue to finance these payments. Thus, any jobs that such measures create come at the expense of other jobs in the state, because the extra expenditure caused by purchasing preferences can only be obtained by higher taxes on other sectors of the in-state economy.

## Why Interstate Trade Barriers Are so Pervasive

There are two potential explanations for state governments' motives for erecting trade barriers: first, state governments are attempting to exploit other states; second, state governments are attempting to redistribute income within their own states. These two possibilities are not necessarily mutually exclusive, but a determination of which condition predominates will aid in designing policies to prevent individual state protectionism. The evidence indicates that the primary motive is redistribution of income within the state erecting the trade barrier. The explanation for such behavior is based on a model of rent-seeking behavior by state governments. The model is used to examine three potential sources of impediments to trade barriers: the in-state losers in the redistribution process, the other states that have their markets restricted, and the Supreme Court.

### *Theories of Governmental Behavior and Trade Barriers*

In one theory, the median voter model represents state governments as behaving "as if" they are individuals. The government's goal in this model is to maximize the well-being of the average or "decisive" individual.[5] In another set of theories, which involve group behavior and the conflicts between groups, one group may "pay" the government to reward it at the expense of other groups within the state. This rent-seeking behavior is unproductive; it is aimed at redistributing income rather than at creating it (see Buchanan et al. 1980). Rent-seeking behavior is not predicted to occur in median voter theories, but it may occur in group conflict models. The existence of interstate barriers to trade seems indicative of rent-seeking behavior.

[5]See Rose-Ackerman (1981) for an application of this model to interstate behavior.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In all cases of administrative import barriers to interstate trade that have been uncovered, a relatively small special interest group receives protection. Labor restrictions, for example, increase the incomes of selected professionals at the expense of individuals purchasing their services. Similarly, the incomes of farmers, insurance companies, banks, and sellers to state governments are bolstered, while the consumers pay the resulting higher prices. In none of these examples is there a net gain to the state; that is, the state as a whole is worse off, even though some narrow groups gain. Prices for in-state consumers rise, and there is no method to compensate them for the loss. If these trade restrictions were subjected to a vote of the population, the losers would greatly outnumber the gainers and the restrictions would be voted down.[6] The explanation for the existence of the restrictions, therefore, is found in rent seeking rather than in the median voter model. When there are trade barriers, one relatively small cohesive group is able to obtain favorable treatment for itself while a large but dispersed group bears the cost. The burden on any individual member of the losing group may not be large, which helps to explain why small groups can be successful with their rent-seeking behavior. Nonetheless, a small transfer from each member of a large group results in a large transfer to each member of a small group. The total amount of money transferred is large and results in significant distortions within the national economy.[7] Given this analysis, policies to restrict interstate trade barriers must seek to control rent-seeking behavior.

There is one way that a state can attempt to make its residents better off than before: by using export taxes.[8] Export taxes may be levied by a state in the belief that residents of other states will pay the bulk of them, resulting in little burden on current state residents. Export taxes, therefore, are ostensibly consistent with the decisive

[6]This is equivalent to saying that a slate of candidates could run and, presumably, win on a platform of removing interstate trade restrictions. Our electoral process is too complicated to generally allow single-issue candidates (and only one issue can win in each election). This is the source of a major problem with representative democracy; we elect candidates, we do not vote on issues. This problem has prompted students of the political process to study rent seeking, a process wherein office holders can take positions that would be unpopular if their position were well known (see Buchanan et al. 1980). Rent seeking is, in part, a formal representation of special interest politics.

[7]This description of in-state redistribution ignores another major cost: the loss of business for firms or producers located in other states. In either motivation for interstate trade barriers, however, the out-of-state producers are losers. Only a rent-seeking motive explains in-state redistribution.

[8]Export taxes have been implicated most often in the past as the source of economic distortion that arises due to federalism (Walker 1969; Posner 1977).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

voter model of governmental behavior, as they represent an attempt to make the average resident better off at the expense of residents of other states. An example of this tax is the Montana coal tax, which was recently ruled on by the Supreme Court. The Court ruled that this tax did not violate the Commerce Clause. The argument is that this tax, 30 percent of the value of coal, is mainly paid by residents of other states because most of the coal mined in Montana is used in other states. What is important here is that the tax would appear to benefit Montana residents, because they receive additional public services or lower taxes as a result of the revenue from coal taxes.

There are two problems with this reasoning: first, it is not apparent that residents of other states pay the tax; second, this may still indicate rent-seeking behavior. It is very difficult to determine tax incidence, that is, who ultimately pays a given tax; but in order for the Montana tax to fall on residents of other states, several conditions must be fulfilled. First, Montana coal producers would have to raise the price by the full amount of the tax. If the industry is competitive, producers will not be able to do so because their customers will go elsewhere. If Montana coal producers have some monopoly power, however, the price may rise. But even in this case the price will only rise by a relatively small proportion of the total tax in the long run, as coal customers will eventually find alternative sources of supply. Thus, part, if not all, of the tax falls on Montana coal producers (see McLure 1983). This element of the tax may be ultimately paid by coal workers through lower wages, by the landowners through lower royalties, or by the owners or stockholders through reduced profits. These groups are most likely to include Montana residents. The same analysis would apply to other export taxes. Therefore, a more likely explanation for the export tax is that it represents an opportunity for Montana residents to tax a narrow group, the coal producers.

*Public Policies Concerning Trade Barriers*

The preceding analysis strongly indicates that administrative trade barriers, the most common form of restriction, appear to be stimulated by rent-seeking motives. Understanding this motivation is crucial for an analysis of public policies concerning trade barriers. The primary agency for controlling trade barriers is the Supreme Court, which primarily has been concerned with limiting interstate exploitation rather than in-state redistribution. Unfortunately, this corresponds to attempting to restrict the minor motivation for trade barriers and neglects what appears to be the major motivation. Other potential policy actions are also examined, such as retaliation by other state



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

governments. Not surprisingly, however, they are also inadequate for restraining the level of barriers to interstate trade.

The Commerce Clause is the main limitation on barriers to interstate trade, and the enforcement of this clause has fallen to the Supreme Court. The Court, however, generally has decided Commerce Clause cases on very narrow grounds based on the particulars of each case; it has been unable to deter state governments from enacting further restrictions on interstate trade. Thus, the primary source of power that can limit interstate trade barriers has not been effective. There are other forces that work to limit trade barriers, but they are not very strong compared to the forces that operate to encourage further enactment of trade barriers.

The Supreme Court has a difficult job. It must decide whether a particular state policy actually is designed to limit interstate commerce or legitimately is designed to protect the health and welfare of the residents of the state. In general, the Court will allow almost any administrative barrier, as long as it does not appear to be explicitly aimed at an out-of-state producer. Conversely, the Court almost universally has ruled out differentials in taxes between goods produced within or outside the home state.

There are two notable exceptions to the "no tax" rule, both of which are illustrative of the general problems facing policymakers in the area of interstate barriers to trade. The first exception is insurance. In a well-known case, the Supreme Court ruled that trade in insurance is not commerce in goods, so it is not subject to the Commerce Clause. This peculiar rule allowed states to explicitly tax out-of-state insurance companies while exempting firms located within the state. In 1944, the Court overturned its earlier ruling and declared that insurance industry regulation was subject to Commerce Clause restraints. In 1945, however, Congress passed the McCarran-Ferguson Act that restored the regulation of insurance to the pre-1944 standards. Congressional interference in this instance is puzzling, especially in light of the fact that the Court has frequently requested that Congress legislatively set general guidelines on interstate trade issues. Congress has never provided general guidelines concerning interstate trade to the Court, even though it chose to act in this particular instance.

The other exception to the "no tax" rule is the alcoholic beverage industry. The 21st Amendment, which repealed Prohibition, leaves all regulation of alcohol to individual state governments. This has been interpreted to include differential taxation of in-state and out-of-state producers. Many states, for example, protect home-produced wines via a larger tax on out-of-state wines. States also frequently



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

protect small beer breweries located within their boundaries. It is difficult to justify either of these two exceptions to the only general guideline set down by the Supreme Court (the "no tax" rule), but even this guideline fails as a useful tool for addressing the interstate trade problem.

The Supreme Court has declined to set new legal precedents that broaden restrictions on trade barriers, beyond the simple "no tax" rule, without congressional action. Because of the lack of legislative guidance, the Court has seen a continuous stream of cases, many of which are related. From 1792 to 1932, the Court dealt with 604 Commerce Clause cases, about four per year (Gavitt 1932). During the 1970s, it was still deciding about four cases per year.

The case law that has resulted from all this activity is not particularly useful for restricting trade barriers. There is virtually no economic difference between an administrative restriction and an import tax, except that the import tax creates revenue. The price of the good in question still rises, rent seeking in the form of administrative barriers is unimpeded, but there is no corresponding tax revenue to compensate the losers of the redistributive struggle. The Supreme Court has ensured that consumers lose, without safeguarding the economic interests of either the state that imposes the trade barriers, the other states that have their goods discriminated against, or the economic health of the nation. This situation represents a breakdown in the smooth functioning of the federalist system and arises because Congress has not armed the Supreme Court to control rent-seeking behavior.

Other factors that might operate to limit the extent of barriers to interstate trade can be examined in light of the apparent underlying motivations for trade barriers. One possibility is that some states may attempt to limit the extent to which they are exploited by others. But this possibility generally should not be expected to occur. Instead, the primary motivation for erecting interstate trade barriers appears to be in-state redistribution of income; political forces *within* each state, therefore, will primarily determine the extent of trade barriers. And these barriers will produce both winners and losers, so that any federal policy that hopes to limit trade barriers can do so only if it alters the relative strength of the two groups.

Interstate retaliation presumes that the underlying motivation for trade is exploitation of other states. One of the few retaliatory cases occurred in Pennsylvania. The state began to vigorously enforce a law that provided for a total embargo of state purchases from states that discriminated against goods produced in Pennsylvania. This enforcement is credited for forcing several eastern states to repeal



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

their state government purchasing preference laws (Jordan 1978). Such cases, however, are rare because direct costs on outside parties are difficult to identify. Moreover, the harmed out-of-state interest must be of sufficient strength to obtain the aid of its state government.

Policymaking in interstate trade has failed because no policy instrument has been directed toward reducing the forces that motivate the erection of trade barriers. We have attempted to demonstrate that interstate barriers to trade mainly arise out of attempts to redistribute income within the state. Because this is not the motivation to which current policy instruments have been directed, they have not impeded the flow of trade barriers. Neither the natural forces of the injured parties nor federal action is equipped to deal with this problem. A new set of policy machinery or a major redirection of current policy instruments is needed to protect the national economic interest.

## Potential Policies to Limit Interstate Trade Barriers

The effect of the Supreme Court's interpretation of the Commerce Clause has been to limit the ability of state governments to directly exploit the citizens of other states. That was, no doubt, one of the intentions of the framers of the Constitution. What the Supreme Court has failed to do, however, is to protect the general citizens of a state from policies that economically protect narrow groups, that is, from rent-seeking behavior by state governments. Briefly restated, administrative trade barriers that do not benefit a state appear to be motivated by rent seeking, with the result of protecting narrow economic interests at the expense of the general public. An additional, although perhaps unintended, result of interstate trade barriers is that producers in other states now have more restricted markets. In sum, the economic distortions caused by interstate trade barriers consist of higher consumer prices and less total output in the protected state. There is also less production and lower profits in other states. The economic restrictions that result from interstate trade barriers are severe in the aggregate because they are so pervasive. Federal policymaking currently neglects this important source of economic inefficiency.

The primary agency for restoring economic order is the Supreme Court. Perhaps the most important reason the Court has been unable to stop the proliferation of barriers to interstate trade is that it has not had the desire to impose broad restrictions. Such restrictions would reverse 200 years of case-by-case determination. An indication that the Court recognizes that a policy problem exists is in its systematic advocacy for congressional action concerning public policy in this



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

area. The reasons for congressional inaction may well be the same as those for the Supreme Court's refusal to act.

There are several conflicting issues that defy solution by sweeping generalization. In particular, the law itself appears to be in conflict, as was seen in the use of the 21st Amendment versus the Commerce Clause. Even more broadly, however, determining whether state actions harm the national interest or harm the national interest more than the value of permitting independent state action presents a host of conflicting and hard to balance issues.

It is difficult for any administrative body to decide whether a particular state policy has been imposed to protect a special economic interest or if the policy may be justified on other grounds. There are several types of legitimate concerns that should generate latitude for state governments passing laws that differentiate over state lines. Some examples illustrate situations where it has been difficult to determine how much latitude is appropriate for state governments.

In New Hampshire, a commuter tax was imposed on people who work in the state but who reside outside New Hampshire. Many cities have similar taxes for people who reside in the suburbs. When the tax was challenged before the Supreme Court on the grounds that it was distortionary, however, the Court upheld the tax. The Court argued that a tax is not distortionary if the state is compensated for services provided to out-of-state residents. This distinction, however, may be difficult to determine in practice.

In Maryland, the state initiated a bounty program to encourage junk dealers to remove abandoned cars from the streets. In 1974, however, the state changed the documentation required to collect the bounty to prevent junk dealers in West Virginia from participating. There is no reason why Maryland residents should pay to remove junk from West Virginia. One of the important features of our federalist system of government is that it allows states to pursue potentially innovative policies. In this case, the Supreme Court apparently concurred with this view and overturned a lower court ruling to uphold the Maryland law.

Certain agricultural trade restrictions may be justified as necessary to prevent the spread of disease, but this justification could be inappropriately used to protect home-state growers. The Texas law prohibiting Florida grapefruit defined mature grapefruit as containing 9.0 parts sugar to 1.0 part acid; Florida grapefruit usually contains 7.5 parts sugar to 1.0 part acid, making it immature and inedible under Texas law. A Florida newsman reported that "local officials expressed the belief the Texas legislation was passed completely as



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

a protective measure for Texas growers." The law was later repealed without any noticeable health deterioration of the local population.

Illinois passed a law that mandated that trucks operating in the state use a type of mudguard that was different from that customarily used and required by other states. The law imposed a significant burden on interstate truckers, who had to stop at the state line to install the Illinois-specific mudguards. The Supreme Court eventually declared this law unconstitutional (Thorpe 1977). On the other hand, the Court upheld a statute that allows different states to limit the maximum allowable width for trucks, even though this width varies among states. The rationale for such a law is that the width requirements change to account for variation in the conditions for safe operation. It is difficult to differentiate between these two laws, because the second does mean that certain trucks cannot operate in all states.

These examples illustrate why the Supreme Court has been reluctant to pass sweeping judgments; some interstate barriers seem worth the price of economic distortion, and others do not. There are two conflicting issues here: first, federalism allows each state to tailor laws to its particular situation, such as truck widths; second, this possibility may tempt state governments into using federalism to pass laws redistributing in-state income. Determining the consequences resulting from an interstate trade barrier requires a significant degree of economic expertise. The trade-off between the economic gain from allowing local law to vary must be balanced against the economic loss from restricted trade opportunities. Further, while there are many reasons to suspect that state governments themselves do not correctly value these trade-offs, it is certain that states generally do not value costs borne by residents of other states. Neither the Supreme Court nor Congress has the economic expertise to determine the level and distribution of economic costs on a case-by-case basis.

Black (1981) proposed that an administrative division within the Federal Trade Commission or a similar agency be empowered to differentiate economically valid state policies from barriers to interstate trade. The division would adjudicate potential infractions of the Commerce Clause. An infraction would be judged to occur when the costs of the interstate trade barrier outweighed the benefits, irrespective of whether or not those costs accrued to residents of the state. A federal agency would not be subject to the same pressure groups as would a state government, and so it should be relatively immune from arguments to redistribute income within the state. This would appear to remove the major cause of interstate trade barriers.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

There are also some disadvantages of giving oversight to Black's proposed administrative division. It is not clear that administrative agencies can fulfill their stated objectives without causing a new set of problems. A more fundamental problem is that a federal agency that would rule on state laws potentially poses a threat to the delicate balance between state and federal governmental rights. In addition, approval of a barrier by a federal agency for one state might be expected to produce 49 similar barriers in others.[9]

Alternatively, Congress could provide the Supreme Court with stricter guidelines for limiting interstate trade barriers. Despite the difficulties in constructing such guidelines, there is substantial room for improvement over the current situation. It has already been shown that eliminating import taxes without limiting administrative import restrictions ensures that consumers lose the rent-seeking struggle being played out in state legislatures. Guidelines for the Supreme Court that would also eliminate administrative import barriers would provide substantial aid to the free functioning of our economy. For example, Congress could require that any trade restriction with an equivalent effect on interstate trade as a tax on goods produced out of state should be treated like a tax and declared unconstitutional.

The real question is whether or not the current system is fulfilling the needs of the country. No estimate of the damage done to the economy because of interstate trade barriers is available. What evidence there is, however, indicates that the damage is substantial, one "guesstimate" being $397 billion annually in lost output. Economic losses approaching this magnitude would appear to be sufficient to inspire the federal government to take some policy steps to restore the lost output resulting from this self-destructive behavior.

## Conclusion

The federalist system of government has served the nation relatively well, allowing local tastes to be expressed independently of the federal government. The system does have drawbacks, however, one of which is that states will be tempted to act as if they are separate national governments; that is, they will enact policies designed to enrich their own residents at the expense of residents of other states. Even worse, state governments may use out-of-state residents as pawns in a struggle to redistribute income within the state. Barriers

[9]This appears already to be the case. As soon as one state-imposed trade barrier passes a Supreme Court test, other states are ready to imitate the barrier. The purpose of the agency would be to erect one further step and, therefore, raise the cost of erecting these destructive trade barriers.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

to interstate trade, where in-state income redistribution appears to be the major motivating factor, impose substantial costs on the economy through reduced opportunities to trade for the best available goods. These are the type of costs the framers of the Constitution attempted to prevent through the Commerce Clause. After 200 years, perhaps it is not surprising that this portion of the governmental system is in need of repair.

## References

Benham, Lee, and Benham, Alexandra, "Regulating Through the Professions: A Perspective on Information Control." *Journal of Law and Economics* 18 (October 1975): 421–48.

Benson, Paul R., Jr. *The Supreme Court and the Commerce Clause, 1937–1970.* New York: Dunellen, 1970.

Black, Charles L., Jr. "Perspectives on the American Common Market." In *Regulation, Federalism, and Interstate Commerce,* pp. 59–66. Edited by A. Dan Tarlock. Cambridge: Oelgeschlager, Gun, and Hain, 1981.

Buchanan, James M.; Tollison, Robert; and Tullock, Gordon. *Toward a Theory of the Rent Seeking Society.* College Station: Texas A&M Press, 1980.

Council of Economic Advisors. *Economic Report of the President.* Washington, D.C.: Government Printing Office, 1985.

Craig, Steven G., and Sailors, Joel W. "Interstate Trade Barriers in a Federalist Economy." Center for Public Policy Discussion Paper 84–5, 1984.

Gavitt, Bernard C. *The Commerce Clause of the U.S. Constitution.* Bloomington, Ind.: Principia Press, 1932.

Greene, Karen. "Occupational Licensing and the Supply of Nonprofessional Manpower." Manpower Research Monograph No. 11, U.S. Department of Labor, 1969.

Hillman, J. S., and Rowell, J. D. "A Summary of Laws Relating to the Interstate Movement of Agricultural Products to the Eleven Western States." Tucson: University of Arizona Agricultural Station, 1952.

Jordon, Donald E. "In-State Preferences in Public Contracting: States' Rights versus Economic Sectionalism." *University of Colorado Law Review* 49 (Winter 1978): 205–31.

Magee, Stephen P. "The Welfare Effects of Restrictions on U.S. Trade." *Brookings Papers on Economic Activity* (1972): 645–701.

Massel, Mark S. *Competition and Monopoly: Legal and Economic Issues.* Washington, D.C.: Brookings Institution, 1962.

McLure, Charles E. "Tax Exporting and the Commerce Clause." In *Fiscal Federalism and the Taxation of Natural Resources,* pp. 169–92. Edited by Charles McLure and Peter Mieszkowski. Lexington, Mass.: D.C. Heath, 1983.

Melder, Frederick E. *Public Purchase Preference Laws as Trade Barriers.* Trade Barrier Research Bulletin Series, The Council of State Governments, n.d.

Posner, Richard A. *Economic Analysis of Law.* 2nd ed. Boston: Little, Brown, 1977.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Rodgers, John M. *State Estimates of Interregional Commodity Trade, 1963.* Lexington, Mass: D.C. Heath, 1973.

Rose-Ackerman, Susan. "Does Federalism Matter? Political Choice in a Federal Republic." *Journal of Political Economy* 89 (February 1981): 152–65.

Shepard, Lawrence. "Licensing Restrictions and the Cost of Dental Care." *Journal of Law and Economics* 21 (April 1978): 187–202.

Thorpe, William L. "Constitutional Law—Commerce Clause-State Purchasing Activity Excluded from Commerce Clause Reveiw—Hughes vs. Alexandria Scrap Corp." *Boston College Industrial and Commercial Law Review* 18 (June 1977): 893–928.

Walker, J. L. "The Diffusion of Innovations Among the American States." *American Political Science Review* 63 (September 1969): 880–99.

Whitehead, David D., and Luyties, Jan. "Can Interstate Banking Increase Competitive Market Performance? An Empirical Test." Federal Reserve Bank of Atlanta *Economic Review* (January 1984): 4–10.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# West Germany's Economic Policy: What Direction?

*Peter Trapp*

## The Revival of the Locomotive Strategy

Recently there has been a revival of the "locomotive strategy," which was first applied in 1977–78. At that time, politicians and many economists argued that a lack of demand could stall economic recovery, that unemployment would remain on an extremely high level, and that serious external imbalances would persist (Gebert and Scheide 1980). As a result, "strong" countries—those whose economies had a low rate of inflation and a current account surplus—were pushed to adopt more expansionary monetary and fiscal policies. The philosophy behind the strategy was that the strong countries, namely, West Germany and Japan, had a special responsibility for the world economy. Proponents of the locomotive strategy argued that because these economies were in a healthy position they could alleviate the problems of "weak" countries, which otherwise would be forced either to adopt unduly restrictive policies or to take recourse to protectionism.

When the upswing in the economy gained momentum in 1979, the locomotive strategy was followed by the "convoy approach," which directed strong economies to guide weak ones toward more growth and less unemployment. On the way, however, the convoy steered into an inflationary storm and the whole fleet fell apart.

Especially in the strong countries, inflation rates increased markedly and the current account moved into deficit. In West Germany, for example, inflation accelerated from about 1 percent in the second half of 1978 to almost 8 percent in the first half of 1981, and the current account moved from a surplus of about 1.5 percent of GDP

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at the Kiel Institute. An earlier version of this paper was presented at the Cato Institute Policy Forum, May 21, 1986.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

(18 billion DM) in 1978 to a deficit of 2 percent of GDP (28 billion DM) in 1980. After that experience, the locomotive strategy seemed to lose ground.

During the early 1980s, there was a broad international agreement that inflation rates had to be brought down and that employment and growth had to be strengthened, but not by stimulating demand. Instead, improvement of the basic conditions for growth and employment was emphasized. Supply-side policy was the new recipe. There was no consensus among governments, however, on what kind of measures had to be taken. In fact, very different policies were pursued under the name of supply-side policy.

In the United States, a restrictive monetary policy aimed at reducing inflation was combined with a significant (by European standards) tax cut and a marked expansion of public expenditures. Some markets were deregulated (trucking, airlines, banking, and energy) in order to increase competition. Another important feature that distinguished the United States from European economies was the flexibility of the labor market in terms of labor cost and manpower (OECD 1985, p. 104).

In West Germany, policymakers implemented a tight monetary and fiscal policy; the reduction of inflation and of the external and internal (budget) deficit were considered to be the preconditions for a sound economic recovery. In order to narrow the public sector deficit, taxes were increased and expenditure growth was kept below the growth rate of nominal GNP. Some cases of income transfers to private households were cut slightly, but most of these steps were later reversed. The increase in labor costs moderated only marginally.

In France, the socialist government that came into office in 1981 did quite the opposite. Monetary targets were moved upward and public expenditures were expanded strongly. In particular, welfare payments were extended and minimum wage earners were granted above-average wage increases. It was hoped that the resulting demand push would ignite an even faster growth of investment, spearheaded by the expansion of the nationalized industries, and that this would ultimately result in less unemployment, less inflation (because of the supply effect of investments), and a self-financing of the expenditure program.

Thus, after having tried to establish a close coordination of macroeconomic policies under the locomotive strategy during the late 1970s, there was a period of policy competition among major industrialized countries during the early 1980s. To some extent, the authorities used the scope for autonomous policies that is given under a system of flexible exchange rates (Vaubel 1980). But soon it became



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

clear that autonomy did not mean that the country could insulate itself from economic influences from abroad. Freedom had its price.

The French experiment failed rather soon, because the exchange rate arrangements within the European Monetary System (EMS) had reduced the scope for autonomous action. Because the Mitterand government preferred to defend the franc rather than to stick to its program of demand expansion and social reforms, it reversed its policy stance in 1983 and adopted a strategy more or less in line with West Germany's.

The West German government achieved its goals. The rate of inflation declined, the budget deficit was reduced substantially, and the current account moved into surplus in 1982. In spite of these improvements in the underlying conditions of economic growth, however, the increase of GNP since 1983 was noticeably weaker than in previous upswing periods and unemployment continued to rise until mid 1985.[1]

The U.S. approach proved to be more effective, particularly in terms of growth and employment. From 1983 to 1985, real GNP rose at an average annual rate of 4.2 percent compared to only 2.5 percent in West Germany and 1.2 percent in France. Even more remarkable was the expansion of employment, which contrasted sharply to the poor employment performance in Western Europe. The growth differential between the United States and European economies was hardly sufficient to explain the difference. It is probable, however, that greater flexibility in the U.S. labor market and economic deregulation contributed greatly to the pronounced improvement in employment. The rate of inflation, which had declined from 13.6 percent in 1980 to 3.2 percent in 1983, remained moderate despite the strong upswing in 1983–84. No progress was made, however, with respect to the federal budget deficit. It had not been reduced by the upswing; on the contrary, it widened significantly, as did the current account deficit.

Notwithstanding the twin deficit, foreign investors were attracted during the first half of the 1980s by the new strength of the U.S. economy. The result was a strong appreciation of the dollar. For companies, it would have been quite normal (and necessary) to imitate a more successful competitor in order to catch up to and defend market shares. The European governments, however, considered the flight from their currencies into the dollar as an undue interference in domestic policies. In order to prevent higher import prices from

[1]From 1983 to 1985, the average annual growth rate of GNP was 2.5 percent compared to 4 percent after the recession of 1974–75 and 6.5 percent after the 1967 recession.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

fueling domestic inflation, they felt forced to raise interest rates. High interest rates, however, were seen as the major impediment to the upswing; therefore, they called for a reduction of the U.S. budget deficit instead of improving growth conditions by lowering marginal tax rates and deregulating markets in their own economies.

The different strategies predictably led to different outcomes and to strong exchange rate reactions. From 1980 to the spring of 1985, the dollar appreciated markedly vis-à-vis the European and the Japanese currencies. U.S. authorities showed little concern about the strong dollar and the high deficits as long as the economy advanced rapidly. But during 1984, economic growth weakened and the Reagan administration began to worry about the loss of competitiveness due to the strong dollar. In 1985, when contractionary effects were anticipated from the implementation of Gramm-Rudman-Hollings, Secretary of the Treasury Baker and others rediscovered the urgent need of international cooperation in the field of demand management in order to harmonize the results of economic policy in industrial countries. As in 1977–78, countries with a low rate of inflation and a current account surplus, such as West Germany and Japan, were singled out for more expansion.

Recent U.S. policy recommendations differ little from the advice Robert Solomon offered the 1978 Bonn Summit; namely, that by adopting the locomotive strategy

> the risks of exacerbating inflation would be minimal. But real benefits would result, primarily for the industrial countries themselves in the form of higher income and lower unemployment. The threat of creeping protectionism would be blunted, investment would be encouraged, and a better pattern of balance-of-payments positions would emerge. In particular, the very large surpluses of Germany and Japan would be reduced.

Likewise, in 1986, Anthony Solomon argued:

> Now that exchange rates are on a more reasonable course and oil prices are declining, there is every likelihood the inflation rate in Germany will be negative in 1986. Clearly, that ought to offer immense opportunity for the German authorities to provide stimulus to lower the current-account surplus, to lower unemployment and to contribute to a better balance in the world economy without threatening any outbreak of inflation.

Are German authorities sacrificing a good deal of employment and output that could be mobilized by "feeding some marks to its economy and advancing its tax cuts"? Or do low growth and the employment problem in West Germany and elsewhere in Europe stem from other sources so that they cannot be solved by just expanding demand?



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

To answer these questions it is useful to look at recent economic developments in West Germany.

## Economic Growth in West Germany Since 1983

Overall output in West Germany has gradually recovered since the trough of economic activity at the end of 1982. In the final quarter of 1986, real GNP was about 12 percent higher than in 1982; on a per capita basis the increase is in the order of 13 percent. The recovery was initiated by a strong acceleration of monetary growth and by some temporary fiscal incentives to investment. These measures led to an upswing of domestic expenditures in 1983 (Table 1). Monetary expansion was slowed markedly in the second half of 1983, however, when the D-Mark again started to depreciate against the dollar and higher import prices threatened to push up the rate of inflation. Therefore, domestic expenditures more or less stagnated in 1984. From the fourth quarter of 1983 until the first quarter of 1985 the increase in real GNP was almost exclusively due to the increase in net exports.

After the deceleration of monetary expansion had come to an end in 1984, the domestic economy began to recover again. During 1985 and early 1986, it received strong impulses: Monetary policy became increasingly expansionary since mid 1985; at the beginning of 1986, income taxes were cut by about 11.5 billion DM (0.6 percent of GNP);

*TABLE 1*

CONTRIBUTION OF DOMESTIC EXPENDITURES AND NET EXPORTS TO GNP GROWTH IN WEST GERMANY, 1982–86

| | GNP (Fourth quarter over fourth quarter, percentage change) | Domestic Expenditure (Change in percent of GNP) | Net Exports (Change in percent of GNP) |
|---|---|---|---|
| 1982/81 | −1.4 | −1.2 | −0.2 |
| 1983/82 | 3.5 | 4.8 | −1.3 |
| 1984/83 | 2.8 | 0.3 | 2.5 |
| 1985/84 | 2.2 | 2.8 | −0.7 |
| 1986/85 | 2.5 | 4.2 | −1.6 |

SOURCE: Deutsche Bundesbank.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and the drop in oil prices provided a considerable real income gain for the economy.

In spite of these stimulating influences, growth slowed during the final quarter of 1985, and there was even a decline of total output at the beginning of 1986. This was partly due to the lag in the effect of monetary policy and to extremely severe winter weather. In addition, companies and consumers reacted cautiously to the strong appreciation of the D-Mark and the drop in oil prices. Because companies expected a further decline of import prices, production activity was cut back for a while in order to run down stocks of finished goods that had been produced at higher input costs. Consumers hesitated to increase spending because they were uncertain whether or not the increase in purchasing power resulting from the oil price decline would last for a longer period.

During the second quarter of 1986, however, demand and production began to pick up again. In particular, investment in machinery and equipment increased strongly. As capacity utilization in the overall economy was nearly as high as in 1979–80, the peak of the previous business cycle, and as profit expectations had risen noticeably, capital widening became more important. The share of such investment, however, is still comparatively low. Private consumption also rose appreciably. Compared to the same period in 1985, total real expenditures on consumption and investment were some 4 percent higher during the second half of 1986. Although the upswing is already more than four years old, the economy is now close to price stability. In the summer of 1986, the index of consumer prices was even somewhat lower than 12 months earlier because of the sharp decline in import prices (−20 percent on a year-over-year basis).

The labor market has improved somewhat. Since reaching a low point in 1984, employment has increased by some 600,000 and short-time work has declined by half a million. The number of unemployed remained high because the labor supply was growing due to demographic reasons and a higher participation rate. In the course of 1986, however, unemployment went down slightly.

What are the lessons to be drawn from the experience of the years 1983 to 1986? First, despite the pronounced reduction of the public sector deficit and the rate of inflation, total output increased by about 3 percent per annum. The most important force behind this was the increase in domestic expenditures.

Second, there is no evidence that economic growth in West Germany relied excessively on exports (Shilling 1985; de Vries 1986). Although exports grew faster than total output, the upswing got only



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

temporary support from exports. For the period as a whole, the increase in imports was much higher.

Third, the expansionary policies at the beginning of the upswing had only a short-run effect. Because the increase of M1 of more than 12 percent was inconsistent with keeping the inflation rate down, domestic demand began to stagnate again as soon as the policy stimulus had to be withdrawn. Thus, the attempt to increase the growth of overall output in the medium run by temporary stimulatory measures failed.

Fourth, domestic economic activity recovered quickly when monetary and fiscal policy were switched (measured in demand terms) from a restrictive to a more or less neutral policy stance.

Finally, the labor market reacted to the recovery, but not enough to reduce unemployment significantly. The elasticity of employment with respect to capacity utilization was about as high as it was during the late 1970s. The fact that the capacity utilization is already very high suggests that the increase of employment is not hampered by a lack of demand. Instead, the impediments are to be sought in the area of labor costs and insufficient incentives to invest in new capacities.

## Short-Term Growth Prospects of West Germany

In 1987, domestic demand is likely to be stimulated by strong impulses and monetary policy will remain expansionary. The target set for the increase in central bank money is a growth of between 3 and 6 percent. In view of the decline of inflation, the target is quite generous; at any rate, it is too high to keep the rate of inflation at zero. In fact, central bank money has even increased at a current rate on an annual basis of some 8 percent. Narrow money, M1, which has been a reliable leading indicator for the direction of economic activity, expanded at an annual rate of more than 10 percent in 1986.

Considering the strong pressure on the West German central bank to lower interest rates, it seems unlikely that the Bundesbank will take action to restrain the growth of the money supply in order to meet its target. As in 1977–78, it will probably tolerate the high growth of the money supply in order to avoid a further appreciation of the D-Mark. Therefore, domestic demand will presumably continue to expand vigorously during most of 1987. The drastic decline in oil prices combined with the appreciation of the D-Mark will give additional impulses to domestic demand, because part of the real income gain becomes effective only during 1987. In 1987, expenditures on consumption and investment will probably increase by 4.5 percent. While export growth will remain moderate because of the



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

real appreciation of the D-Mark by about 10 percent since the spring of 1985, imports will surge due to buoyant domestic demand.

Since the beginning of 1985, the trade account in constant prices has already deteriorated noticeably (Table 2). This trend is likely to continue; in real terms, the surplus in the foreign balance will decline from 5 percent of GNP in 1985 to about 2 percent in 1987. Real GNP is expected to increase by about 3 percent in 1987.

In 1986, real incomes increased in West Germany; there was a considerable rise of production and employment, and prices remained stable. In 1987, however, inflation is likely to rise again when domestic cost pressures are no longer dampened by falling import prices.

## No Demand Constraints to Growth

On the whole, there is no need for further stimulation of demand in the West German economy. On the contrary, as the economy is already operating close to its capacity limits in many sectors, the expected acceleration of demand is likely to result in some overheating of the economy.

Given the present policy stance and the business outlook, why do some economists insist that German authorities should turn to reflation? One possible answer is that the views on the importance of economic targets differ and that there are different views on the appropriate instruments to achieve those targets. Another answer is that the recommendations are based on an inappropriate analysis of the West German economy. An inspection of the critical comments on West German economic policy suggests that both answers seem to play a role.

*TABLE 2*

WEST GERMANY'S TRADE PERFORMANCE, 1985I–1986IV

| | Ratio of Export Values to Import Values | Ratio of Export Volume to Import Volume |
|---|---|---|
| 1985 I | 100 | 100 |
| II | 102.8 | 101.7 |
| III | 104.1 | 99.3 |
| IV | 104.8 | 99.7 |
| 1986 I | 109.1 | 97.3 |
| II | 111.4 | 94.4 |
| III | 116.4 | 96.3 |
| IV | 114.3 | 94.3 |

SOURCE: Deutsche Bundesbank.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Scope for a more expansionary policy is seen because the rate of inflation has dropped to a very low level. Price level stability, however, has not been achieved by a stability of domestic costs and prices; instead, it has resulted exclusively from the decline in import prices by more than 20 percent. Assuming a strengthening of economic activity in industrial countries, it seems unlikely that oil prices will fall again in 1987; and given the strong increase in the money supply, the D-Mark will probably appreciate only moderately vis-à-vis the dollar over the next year. Therefore, in 1987 price stability will weaken as domestic demand grows stronger. The rate of inflation will then be determined once again by the rate of increase of domestic prices, which will probably be in the order of at least 2.5 percent due to relatively high wage increases and strong demand. The increase of nominal wages in 1986 by more than 4 percent was higher than in 1985, although the inflation rate was noticeably lower than in 1985 and workers benefited from the tax cut. For the 1987 bargaining round, higher inflation rates as well as higher profits give reason to expect wage increases that will be even above those agreed on in 1986.[2] As there will be no improvement in the terms of trade comparable to that in 1986 and as demand is still buoyant, higher wage costs will be translated into higher prices. Therefore, already at the end of 1987 consumer prices will probably increase by more than 3 percent. At present, the growth rate of the money supply is significantly higher than would be consistent with such a rate of inflation. Therefore, a further upturn of inflation rates beyond 1987 is very likely if monetary expansion remains strong.

An increase of the rate of inflation is sometimes regarded as the price that a country has to pay for a higher rate of economic growth. But experience has shown, especially the failure of the locomotive approach during the late 1970s, that there is no positive relationship between inflation and real growth. Economic policy cannot buy more growth (and jobs) with more inflation. Nevertheless, this philosophy is still underlying many comments. By accepting a higher rate of inflation, the United States has not achieved a significantly higher rate of long-run growth than West Germany. Presumably, the favorable performance of U.S. employment during the 1970s and 1980s did not result from tolerating more inflation; rather, it resulted from the greater flexibility in U.S. labor markets. Those who exert pressure on West German policymakers to reflate in order to decrease unemployment seriously ignore an important fact, namely, that the introduction of labor market rigidities and the reduction of wage differ-

[2]In the metal industry wages will rise by 13 percent from 1987 to 1989.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

entials between regions, branches, and qualifications contributed greatly to the increase of the West German unemployment rate during the last decade (Gundlach 1986; Donges 1985).

## Increasing Labor Market Rigidities

An inspection of labor market data since the 1960s shows that there has been a significant deterioration. Until 1973, unemployment rates hovered around 1 percent. During the 1974-75 recession, unemployment rates (number of unemployed as a percentage of the dependent labor force) rose to about 4.5 percent. During the following recovery, from 1976 to 1980, the unemployment rate only declined to 3.5 percent. During the recession of the early 1980s, the rate jumped to about 9 percent. Given present labor market rigidities, unemployment will hardly drop below 8 percent until 1987. In 1985, about one-third of all jobless persons registered at the labor exchange had been unemployed for a year or more, compared to 5 to 10 percent of long-term unemployed during the early 1970s.

What accounts for the high level and the stickiness of the unemployment rate? One reason is that labor costs are comparatively high in West Germany, too high to offer sufficient employment opportunities for less qualified workers or in weak branches or in regions with structural problems. One reason for this is that the wage contracts between employers' associations and unions are based on expected average increases of productivity, profits, and prices. Until the late 1960s, contractual wages were significantly lower than effective wages. During the 1970s, however, the wage drift narrowed and almost disappeared. Because contractual wages are minimum wages for the whole industry, companies with below average productivity gains found it more and more difficult to cope with increased wage costs. German workers have to enjoy not only quite a high income, but they are also entitled to six weeks' annual leave plus 10 to 15 extra holidays (depending on the state they live in), unlimited sick leave, and a work week that is approaching 37 hours. Actual retirement age is well below 60 years. All this adds up to a considerable wage bill. Because early retirement, extensive health care, and relatively generous unemployment benefits are expensive, social security contributions have been raised noticeably. As a share of GNP, they have increased from less than 12 percent during the late 1960s to 17.5 percent during the mid 1980s. Half of these contributions have to be paid by employers; the other half is deducted from the worker's gross income. The latter part has been translated into higher wage increases, because trade unions succeeded in defending net



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

wages. Employers managed to reduce labor costs by cutting employment, however, particularly during recessions and by turning to labor-saving production methods. Nevertheless, total labor costs rose faster than the productivity of labor. Operational profits, that is, profits from productions (not from financial investment), have declined massively over the last 15 years (Dicke and Trapp 1984, 1985). In spite of the recent recovery, these profits are still quite low compared to the early 1970s and only marginally higher than the real rate of return on financial assets.

The high unemployment rate is also affected by workers having less mobility and more wage inflexibility, because the incentives not to work have been increased over the last 10 years. This policy has produced at least three significant developments. First, the marginal tax rate (including social security contributions) is presently at about 60 percent for an average income earner (Boss et al. 1986). Therefore, instead of working overtime many workers prefer to work at home or in the shadow (underground) economy, which is estimated to be 10 percent of the total output of the official economy.

Second, the amount of welfare payments that a family is entitled to when there is no earned income comes close to the net income of an average wage earner (Boss 1986). A family of four without income, for example, can collect welfare payments that add up to at least 78 percent of the average net income of a family of four with one income earner. Because of the high level of welfare payments, additional wage increases for low income groups were agreed upon in 1986 in the public service and in the metal industry.

Finally, the German parliament, the Bundestag, has passed a number of laws since 1970 designed to enhance job safety. Consequently, the costs of layoffs are now four times higher than in 1972 (Soltwedel 1984). In practice, these regulations have made the labor market more rigid and were an important reason for the reluctance of companies to increase employment during upswing periods.

These examples illustrate that the unemployment problem in West Germany is deeply rooted; it is structural in character and not solely attributable to deficient aggregate demand. Therefore, unemployment cannot be overcome by additional demand, as was already evident in 1979–80.

## The Current Account Balance: A Growth Perspective

The advocates of the locomotive strategy view not only unemployment but also the surplus in the current account as indicators of insufficient domestic demand and mentioned them as a reason for



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

reflation. As to the 1986 increase of the current account surplus to some $35 billion (3.5 percent of GNP), it has to be seen that this is not related to the development of real magnitudes but is only the result of a strong improvement of the terms of trade, particularly of the redistribution of the OPEC surplus. Because the West German economy will not experience a similar terms-of-trade improvement in 1987 (more likely there will be a slight deterioration) and because imports will expand rapidly, the surplus in the current account for the year as a whole will probably decrease by 25 percent in 1987 and even more during the course of the year.

Given the high rate of unemployment and insufficient growth of potential output in West Germany, any export of domestic savings, which is the counterpart to the surplus in the current account, may appear to be undesirable from a domestic point of view. In this respect, the adequate remedy would not be to reflate the economy but to make investment more attractive by lowering taxes, deregulating protected sectors, and allowing for more flexibility in the labor market. As a result of such policies on the micro level, the West German current account could move into deficit, which then should be accepted. It would indicate that the West German economy has become more attractive to investors worldwide.

If the U.S. deficit in the current account decreases because the Gramm-Rudman-Hollings act is applied, however, there is no need to stimulate demand in West Germany. When world savings are unchanged and the U.S. current account deficit falls, interest rates could go down further. This would ease the burden for debtor countries and stimulate investment there and in Germany and Japan. On the other hand, if the deficit of the U.S. current account remains high because investment opportunities there are considered to be better than elsewhere, this should give an incentive to other countries to improve investment profitability by cutting taxes, removing subsidies, and deregulating markets.

## Conclusion

It has been stressed that there is no need for more coordination of macroeconomic policies, particularly for a reflationary policy in Germany. But this should be interpreted neither as a rejection of policy coordination in general nor as a general defense of what West German economic policy has done over the last several years. As to coordination, there are many areas where concerted action is urgent, but these do not concern macroeconomic policies; they concern issues such as the world-wide abolishment of regulations and barriers to



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

trade and the elimination of restrictions on capital flows. In particular, the protectionist trade policy pursued on both sides of the Atlantic has to be changed.

As to West German economic policy, the government must do a great deal to live up to the market-oriented policy it promised to implement after the 1983 election. Taxes have to be cut—and by more than is envisaged for 1988 and 1990—in order to ease the tax burden of average income earners and to lower the top rate and the progressivity of the tax schedule. Subsidies amounting to almost 7 percent of GNP must be reduced. A Kiel Institute study shows that a combination of lowering subsidies to 3.5 percent of GNP and of cutting taxes by the same amount over a period of five years could create one million additional jobs (Gerken et al. 1985). Furthermore, the welfare system has to be trimmed and deregulation has to be intensified. Such steps are urgent and would help to revitalize the economy. A further stimulation of demand, however, will lead to the old stop-and-go pattern of economic policy and, therefore, is not to be recommended.

## References

Boss, Alfred. "Finanzpolitik seit 1982: Stärker wachstumsorientiert?" *Die Weltwirtschaft*, no. 1 (June 1986): 47–54.

Boss, Alfred et al. *Nach dem Ölpreisfall—Aufschwung ohne Inflation?* Kiel Discussion Paper no. 118. Kiel: Institut für Weltwirtschaft, 1986.

de Vries, Rimmer. "Germany's Growth Predicament." *World Financial Markets* (June/July 1986): 1–9.

Dicke, Hugo, and Trapp, Peter. *Investment Behavior and Yields in Some West German Industries.* Kiel Working Paper, no. 205. Kiel: Institut für Weltwirtschaft, 1984.

Dicke, Hugo, and Trapp, Peter. "Zur Ertragskraft von öffentlichen und privaten Investitionen." *Die Weltwirtschaft*, no. 1 (June 1985): 70–86.

Donges, Juergen B. "Chronic Unemployment in Western Europe Forever?" *The World Economy* 8 (December 1985): 353–72.

Gebert, Dietmar, and Scheide, Joachim. *Die Lokomotiven-Strategie als wirtschaftliches Konzept.* Kiel: Institut für Weltwirtschaft, 1980.

Gerken, Egbert et al. *Mehr Arbeitsplätze durch Subventionsabbau.* Kiel Discussion Paper no. 113/114. Kiel: Institut für Weltwirtschaft, October 1985.

Gundlach, Erich. "Gibt es genügend Lohndifferenzierung in der Bundesrepublik Deutschland?" *Die Weltwirtschaft*, no. 1 (June 1986): 79–88.

OECD. *United States.* Economic Surveys 1985/86. Paris: OECD, November 1985.

Shilling, Gary. "U.S. Imports: Sailing Through a Narrow Passage." *Business Economics* 20 (October 1985): 18–23.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Solomon, Anthony M. "West Germany Puts Savings over Jobs." *Wall Street Journal,* 26 March 1986.

Solomon, Robert. "Advice for the Bonn Summit." *Journal of Commerce,* 6 June 1978.

Soltwedel, Rüdiger. *Staatliche Interventionen am Arbeitsmarkt: Eine Kritik.* Ph.D. diss., Department of Economics and Social Sciences, University of Kiel, 1984.

Vaubel, Roland. *Internationale Absprachen oder Wettbewerb in der Konjunkturpolitik?* Walter Eucken Institut, Vorträge und Aufsätze 77. Tübingen: Mohr, 1980.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Economic Significance of British Labor Law Reform

*Charles G. Hanson*

## Introduction

Over the past generation the British disease of relatively low economic growth and low labor productivity, accompanied by relatively high inflation and unemployment, has become so obvious that no economist would want to deny its existence. For British economists, and indeed for all British people who take even a superficial interest in economics, international comparisons of living standards have become profoundly embarrassing. As Pollard (1982, p. 2) recently wrote, "The statistics confirm the national consciousness of a staggering relative decline, such as would have been considered utterly unbelievable only a little over thirty years ago." The figures in Table 1 are among many that could be cited to validate Pollard's comment.

Given Britain's abysmal productivity in manufacturing industry by 1980, it is not surprising that the country's average share in world exports of manufactures fell from 20.5 percent during the years 1951–55 to 9.1 percent for the years 1973–77, while over the same period West Germany's share increased from 13.1 to 21.1 percent and Japan's from 4.4 to 14.2 percent (Pollard, p. 12). Thus it became commonplace to talk of the deindustrialization of the British economy, which in the 19th century had been renowned as the workshop of the world.

There is no disagreement about the existence of the productivity problem, but there is fundamental disagreement about the remedy. For example, Brown (1979) argued that the right medicine for the British disease was to persevere with policies that had already been

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

The author is Lecturer in Economics at the University of Newcastle upon Tyne, England. He gratefully acknowledges the helpful comments of Michael Jones-Lee, Bruce Grant, and several other colleagues at the University of Newcastle. Thanks for constructive criticism are also due Charles J. Rowley, Andrew Currie, and Cyril Grunfeld.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 1*

OUTPUT PER EMPLOYED WORKER-YEAR IN 1980
(THOUSANDS OF 1973 U.S. DOLLARS)

| Country | Manufacturing | Services |
|---|---|---|
| USA | 16.8[a] | 14.7[b] |
| Netherlands | 16.0 | 12.1[b] |
| Japan | 14.5 | 10.2[b] |
| West Germany | 13.8 | 12.6 |
| France | 12.8 | 12.8 |
| Belgium | 12.2 | 12.6[b] |
| Italy | 10.4 | 10.9 |
| UK | 6.8 | 8.7 |

[a]Includes mining and quarrying.
[b]1979.
SOURCE: Roy (1982).

tried, including the discredited national plan of 1966, consultations in the National Economic Development Office, and "an incomes policy which really works." In marked contrast were those who felt that in a number of areas, including that of labor relations, some change, not to say radical change, was necessary. Allen (1976) and Pratten (1976) expressed a relatively cautious view of the need for management to devote more time to labor relations to achieve better cooperation between management and the work force and to improve labor productivity. Hayek (1980, p. 64) was more forthright in calling for reform of the law when he wrote, "All I can say with conviction is that, so long as general opinion makes it politically impossible to deprive the trade unions of their coercive powers, an economic recovery of Great Britain is also impossible." In the same vein Minford (1983, p. 7) has called for the elimination of all legal immunities for trade unions. And Rowley (1984, p. 1135) has argued that those immunities were the primary cause of labor turbulence over the past 20 years, which has led to "a relatively low rate of investment in British industry and a correspondingly low rate of productivity growth."

At the general elections of 1979 and 1983 Hayek's condition was fulfilled. On each occasion a Conservative government—pledged to limit trade union coercion—was returned to office, and the pledge was kept in the shape of the Employment Acts of 1980 and 1982 and the Trade Union Act of 1984. But what exactly did these acts provide? Have they led to an improvement in labor relations and labor productivity in Britain? And if so, are these reforms sufficient to bring about the regeneration of the British economy? The purpose of this article is to attempt to answer these questions.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

## Trade Union Immunities and Coercion

The Trade Disputes Act of 1906 gave virtually complete legal immunity from tort to trade unions in all circumstances, and legal immunity to their officials when acting "in contemplation or furtherance of a trade dispute." To all intents and purposes Parliament put trade unions and their officials above the civil law. The great constitutional lawyer A. V. Dicey (1914, p. xlvi) wrote of Section 4 of the 1906 act that "it makes a trade union a privileged body exempted from the ordinary law of the land. No such privileged body has ever before been deliberately created by an English Parliament." This situation was temporarily modified by the Conservative Industrial Relations Act of 1971, but the immunities were reintroduced and even strengthened by the Labour government's Trade Union and Labour Relations Act of 1974 and its Amendment Act of 1976. By 1979 the decision of the House of Lords in the case of *Express Newspapers* v. *McShane* confirmed that trade unions and their officials had the right to spread a strike to parties quite unconnected with the original dispute without fear of redress whenever they thought that such secondary action would help their cause. In addition, where employers had agreed to a closed shop, union officials could force existing employees to join a named union under threat of dismissal without compensation. Thus the coercive powers, which in Hayek's view made an economic recovery impossible, could be (and were) used either against employers or employees, both union members and nonmembers. Members could be compelled to do the union's bidding, while nonmembers could be forced to join.

### *The Closed Shop*

The power of union officials over their members is derived partly from the closed shop. Between 1964 and 1978 the number of employees in Britain covered by the closed shop grew from at least 3.8 million to at least 5.2 million. Following the introduction in 1971 of the right not to be dismissed except in certain clearly defined circumstances, the Trade Union and Labour Relations Act of 1974 and its Amendment Act of 1976 provided that dismissal for nonmembership of a designated union was fair whenever a closed shop (technically a union membership agreement) existed, unless a genuine religious objection could be proved. Faced with the choice between joining a union or dismissal in the period 1975–80, most employees chose to join a union; but hundreds refused and were duly dismissed without compensation. In the absence of legal redress in Britain, three former employees of British Rail took their case to the European Court at Strasbourg, and in 1981 the court ruled that their dismissal contra-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

vened Article 11 of the European Convention on Human Rights and that the British government was liable for substantial costs and compensation.

Responsibility for the spread of the closed shop and its abuse is usually pinned on the trade unions, but it needs to be remembered that a closed shop cannot be introduced or enforced without the agreement of the employer. During the 1970s many employers followed the example of British Rail and agreed to make union membership compulsory without any reference to the views of their employees, arguing that it was convenient to negotiate with union officials who could (supposedly) speak for the entire work force. This situation has been transformed by the government elected in 1979, so that in 1985 Employment Minister Peter Bottomley was able to tell Parliament, "The monopolistic hold of the closed shop has been broken." How was this effected? And is the minister's claim justified?

As mentioned above, in 1979 dismissal for nonmembership in a union where a closed shop existed was always fair unless a genuine religious objection could be proved. That situation has now been reversed by the relevant sections of the Employment Acts of 1980 and 1982; dismissal for nonmembership is now always *unfair* unless a closed shop has been approved in a secret ballot among the relevant employees by a majority of either 80 percent of those entitled to vote or 85 percent of those who voted. Of course it is difficult to achieve such a majority. But there is another, more significant reason why so few successful elections have been held. Since 1980 the official policy of the Trades Union Congress (TUC), to which all the major unions are affiliated, has been to boycott secret ballots of every kind which are required by the new legislation. Consequently only a handful of elections have been held and only a few thousand employees are now formally covered by agreements stipulating that dismissal for nonmembership of a trade union is fair.

This development was highlighted by the Bristol Tugboatmen's case, in which an industrial tribunal at Bristol in May 1985 unanimously decided that three tugmen were unfairly dismissed by Cory King Towage Ltd. for not being members of the Transport and General Workers Union and that the compensation to which they were entitled would have to be paid by the union. It seems that this will amount to at least £20,000 each, and the implications of this expensive lesson will not be lost on the Transport and General Workers and other major British unions. In the future they will be much more wary about putting pressure on an employer, through a threatened or actual strike, to dismiss nonunionists. A trade union does not enjoy paying substantial compensation to a deliberate nonmember.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

But will not many companies continue to enforce the closed shop informally? Can the change in the law break down a custom which in some industries—for example, printing—has been effective for 50 years or more? Certainly life will still be difficult, not to say very unpleasant, for the nonunionist in some occupations. But the evidence suggests that these will quickly become a small and dwindling minority. The biggest closed shops in 1980 were in the state sector and covered all who worked in the railway, water, electricity, gas, and coal mining industries as well as the Post Office and British Telecom—a total of well over a million employees. Early in 1985 the boards of these corporations informed the relevant unions that the closed shop had become unenforceable, and in June 1985 the board of British Rail made it plain that it was no longer a party to the union membership agreement. In other words, union membership is once again a matter of personal choice for every employee of British Rail. Such a complete reversal of policy on this issue by a major employer indicates the effectiveness of the carefully thought out reforms of the Employment Act of 1982, backed as they are by substantial compensation for infringement of the law. As far as the closed shop is concerned, events have disproved the argument that the custom and practice of labor relations cannot be affected by a change in the law. But is the same true of legal immunities for trade unions? This is the fundamental issue.

*Trade Union Legal Immunities*

As has already been described, in 1979 following the decision in the case of *Express Newspapers* v. *McShane* the trade unions and their officials were virtually untouchable in a court of civil law. That situation has been radically changed by the legislation of 1980–84. The Employment Act of 1980 curbed union immunities in two ways: immunity was removed from those organizing or taking part in secondary picketing; and immunity was removed from those organizing certain secondary industrial action, including strikes.

Neither of these reforms affected the immunities of trade unions in any way. Rather, they were directed at pickets and union officials, and therefore any damages awarded by the courts for illegal action were likely to be negligible. The failure of the Industrial Relations Act of 1971 had led many members of the government, including Prime Minister Thatcher, to favor a step-by-step approach to trade union law reform, and in 1979 it was widely thought that if the law were changed overnight to allow employers to take trade unions to court it could not be enforced. The government strategy was wisely based on allowing each change to take root before attempting more radical measures. Following this pattern the Employment Act of 1982



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

made it possible for an employer to seek damages up to a limit of £10,000 from a small union to £250,000 from a large one. However, most employers are mainly concerned with carrying on their business, so they will normally seek an interlocutory injunction against a union (that is, an order to desist from picketing or strike action) if it is acting unlawfully. If an injunction is granted and the union defies it, it becomes guilty of the serious offense of contempt of court, and there is no limit to the fines that may be imposed.

At the same time the act narrowed the definition of a "trade dispute" to disputes between workers and their own employer and made it plain that politically motivated strikes were excluded. This meant that union officials had to consider carefully the motive for their action. If it were politically inspired, they could find themselves and their unions in serious legal trouble. Unions still had legal immunity when their members were engaged in a strike against their own employer for better wages or conditions of work. But if the strike involved unlawful (secondary) picketing, or if it were far removed from the original dispute or being used to impose union membership or recognition requirements on another employer, then an employer could seek an injunction and/or damages against the union.

The Trade Union Act of 1984 took one further significant step to curb trade union immunities. Section 10 provides that immunities are only available to unions that have held a secret ballot among those members who will be called on to strike or to take other industrial action and have received a majority vote in favor of a strike or other action. In other words, the power to decide in favor of industrial action has been transferred from union officials to the members concerned. Although numerous private opinion polls have regularly shown a large majority of trade union members to be in favor of this reform, the official policy of the TUC has been one of total opposition. But by June 1985 this antidemocracy policy was beginning to crumble as a number of unions, including the National Union of Railwaymen, which had previously opposed secret ballots among the members, decided to accept prestrike ballots as a normal arrangement. This change of policy may have been partly prompted by the defeat of the National Union of Mineworkers (NUM) in March 1985, after the senior officials of that union had denied their members the opportunity to vote, although the union's rules provided for a ballot before a national strike.

## Use of the Reformed Law, 1983–85

British trade unions and their officials still have certain legal immunities, but they are conditional on a majority vote in favor of a



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

strike in a prestrike ballot; also, they are only available where there is a genuine trade dispute and a strike is conducted in accordance with the criteria laid down in the Employment Acts of 1980 and 1982. But are employers using the new legislation, and if so, is it proving effective? The short answer to these questions is "yes." The short answer, however, is not an adequate answer. It is necessary to consider some of the instances between 1983 and 1985 in which the new laws have been used more or less successfully and one instance in which the employer could have used the legislation and deliberately refrained from doing so.[1]

*Messenger Newspapers Ltd.* v. *National Graphical Association*

The first major dispute to test the restriction on trade union immunities introduced by the Employment Act of 1982 occurred at Warrington, Lancashire, late in 1983. Eddy Shah, chairman of the Stockport Messenger Group, a chain providing free weekly newspapers, decided to employ some nonunion labor at his printing plant. This move was vigorously opposed by the National Graphical Association (NGA), a union that is dedicated to the closed shop and includes practically all of the skilled print workers in Britain. Shah dismissed six employees (all members of the NGA) who opposed his new policy, and they were quickly joined on the picket lines outside his factory by many other NGA members determined to prevent the printing and distribution of Shah's newspapers. This action, deliberately organized by the union, was in open defiance of Section 16 of the Employment Act of 1980, which specifically limited legal picketing to an employee's own place of work.

Shah duly sought injunctions against the NGA to stop the mass picketing and to prevent the union from bringing pressure on advertisers to withdraw from the Messenger papers. Both injunctions were granted on November 14, 1983, but the unlawful action continued and on November 17 the union was fined £50,000 for contempt of the injunctions by the Manchester Crown Court. On November 22 the NGA national council decided not to pay the fine, and mass picketing with occasional violence continued at the Warrington plant.

Three days later the High Court fined the NGA a further £100,000 and ordered the sequestration (confiscation by court officials) of £175,000 from the union. On the same day the NGA decided on a two-day stoppage in Fleet Street and the Newspaper Publishers

[1]The main sources of information for the industrial disputes referred to in the remainder of this section were press reports in the *Financial Times*, *The Times*, and the *Daily Telegraph;* and the chronicle of the *British Journal of Industrial Relations* from March 1984 to March 1985.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Association threatened to sue for damages. Subsequently seven national newspapers were granted a High Court injunction banning a repetition of the stoppage that had prevented publication of national papers on November 26 and 27.

Talks took place between Shah and the NGA, but these came to a halt on December 9 when the union was fined a further £525,000 in the High Court. The NGA threatened a further one-day national newspaper strike, but this was subsequently called off in the face of injunctions sought by several newspapers. On January 18, 1984, the union's national council decided to accept defeat and to purge its contempt of court in order to unfreeze its assets (of about £10 million), which had been frozen since November. The council's decision proved that the new legislation could be completely effective. The implications of this case are profound because newspaper printing has long been one of the most backward and unproductive industries in the British economy, partly because of the resistance of the NGA and its fellow union SOGAT (Society of Graphical and Allied Trades) to the introduction of new technology and sensible manning arrangements. These unions have long been regarded as among the toughest and most restrictive in Britain. Very soon after his legal victory Mr. Shah began to plan for the publication of a new national daily newspaper using all the latest technology, and it is widely believed that he was the catalyst who brought about the long overdue modernization of the national newspaper industry.

As early as 1962 it became well known that Britain's national newspapers, which regularly denounced other firms and industries for poor productivity, were themselves seriously overmanned. Using a firm of management consultants, the Royal Commission on the Press—which published its final report in 1962—established that four national daily newspapers had 53 percent more staff on the production and distribution side than was required. Criticism was made of trade union insistence on overmanning, but this was followed by the remark that "the ultimate responsibility for all excess costs must rest with management." Moreover, a specific comment was made about "the lack of a really senior Management Representative actually present and in control when the paper is produced" (i.e., between 6 p.m. and midnight). Little or nothing was done about this scandal until the imminent appearance of Mr. Shah on the national newspaper scene. The full extent of the scandal was revealed only in January 1986. When the 5,300 production and ancillary workers producing four national newspapers in central London for News International went on strike, the proprietor, Mr. Rupert Murdoch, sacked them all. He then hired a replacement labor force of 1,300 to



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

produce the same newspapers in a new technology plant at Wapping, East London.

The ensuing dispute between News International and SOGAT, the main union involved, has been extremely bitter both at Wapping, where it quickly became clear that the high fences and barbed wire surrounding the new plant were essential to keep the pickets at bay, and in the courts, where SOGAT's entire financial assets of many millions of pounds were sequestrated for its defiance of a High Court order to stop blocking the wholesale distribution of the company's papers. In June 1986 Mr. Murdoch offered the unions most of his Grays Inn Road print plant and £50 million to settle the dispute, but this offer was rejected in a ballot by the dismissed print workers, so that the dispute was still unresolved in September.

Meanwhile, Mr. Shah's new national daily paper *Today* was launched in March 1986, using all new technology, and other national newspapers were modernizing apace, including the *Daily Mirror* and the *Daily Express*, where the excessive labor forces were substantially pruned. For example, at the *Daily Express* between November 1985 and September 1986, a 38 percent reduction in the work force was negotiated. The 2,500 redundancies cost about £65 million but should produce annual savings of £50 million. Thus it could safely be said in 1986 that after a generation or more of very low productivity, Britian's national newspapers were at last being manned in a more sensible way.

*The National Coal Strike, March 1984–March 1985*

The ostensible cause of the national coal strike, which lasted from March 1984 to March 1985, was the proposal by the National Coal Board (NCB) to reduce output over the financial year 1984–85 by 4 million tons, with a loss of 20,000 jobs. From the beginning of the strike several areas, including Nottinghamshire, continued to work after a ballot had shown a large majority of these men opposed to a strike. There was illegal mass picketing by miners from Yorkshire and elsewhere, and on March 14 the High Court granted an injunction to the NCB ordering the Yorkshire area of the National Union of Mineworkers to stop organizing flying pickets and to withdraw instructions to its members to picket. The picketing continued but on March 19 the NCB sought and was granted an indefinite adjournment of its contempt of court application against the Yorkshire area of the NUM. Thus throughout the lengthy dispute the NCB was unwilling to follow Eddy Shah's example and to use the new legislation against illegal mass picketing. The reason for this policy, and for the unwillingness of the government to instruct Ian McGregor,



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

chairman of the NCB, to use the civil law against the NUM, has still not been explained.

However, legal action was taken against the NUM, both under the new legislation and under the common law. In April 1984 Read Transport Ltd. secured an injunction against the NUM South Wales area to stop secondary picketing outside its depot on the grounds that the picketing was unlawful under the Trade Union and Labour Relations Act of 1974, as amended by the 1980 act, and that the union, through its agents, was committing the tort of inducing Read to break its commercial contracts. The union continued the picketing and in July the High Court imposed a fine of £50,000 for contempt and a subsequent order for seizure of the funds. By mid-August the sequestrators had taken control of £707,000 and the union backed down. The many other private road haulers used by the British Steel Corporation (BSC) and the National Coal Board (to take coal to power stations) did not thereafter find any need to go to court. On the other hand, the big public sector organizations (BSC, NCB, and British Rail) who at various times had good legal grounds for an injunction for unlawful picketing and boycotts did not pursue their rights.

Also, in August two Yorkshire miners, Messrs. Taylor and Foulstone, brought a High Court action against the NUM claiming that the strike was unofficial because no prestrike election had been held and that the democratic processes in their union had become "totally paralyzed." In September Justice Nichols ruled that the Yorkshire area strike should not be called "official," and in a separate case brought by three Derbyshire miners he ruled that the strike in that area was unlawful. A month later, when the union had taken no notice of his rulings, he fined the NUM £200,000 and the president of the union £1,000. The president's fine was paid anonymously, but the High Court ordered the seizure of all NUM national assets after the union had failed to pay its fine. The sequestrators found themselves engaged in lengthy negotiations with banks in Dublin and Luxembourg, as the president had anticipated legal action against the NUM and had secretly moved most of the funds abroad.

In the case of this strike, legal action was not obviously a main cause of its conclusion. By January 1985 the trickle of striking miners returning to work became a stream, and by February the stream had become a flood. In March the NUM called off the strike and claimed victory for the union, while everyone else saw the return to work as a sign of defeat.

*The Three Disputes of March–May 1985*

Section 10 of the Trade Union Act of 1984, which removes legal immunities from unions when a prestrike election has not shown a



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

majority in favor of a strike, was tested three times in quick succession between late March and mid-May 1985 and was found to be exceedingly effective.

Early in March 1985 the national executive of the Civil and Public Services Association (CPSA), a major civil service union, threatened a one-day strike on April 1, and further action if necessary, in support of their pay claim. Following the threat, the Treasury (as the employer of the civil servants) applied for a High Court order against the union on March 25, and on the following day the executive of the CPSA decided to poll its membership about the proposed strike. Three days later the government raised the pay offer from 4 percent to 4.4 percent, and on April 11 the result of the ballot was disclosed. Members of the union had voted against strike action by a majority of 460 out of 90,600 votes cast. On April 23 the pay offer was raised to 4.9 percent and by then all talk of strike action by the CPSA had ceased.

At almost the same time as the CPSA dispute it seemed as though the national postal service might come to a halt. By March 1985 the Post Office was pressing vigorously for substantial changes in working practices, including the introduction of new letter-sorting equipment and the recruitment of many part-time employees. On March 13 it seemed that deadlock had been reached when the Union of Communication Workers informed the Post Office that key sections of the package of reforms of working practices would not be accepted, and the Post Office stressed that, regardless of the agreement, the reforms would be implemented in April and would boost individual income by 5–10 percent. When further talks broke down, 800 of the staff at the Mount Pleasant sorting office in London (the largest in Britain) walked out on April 2 in defiance of a High Court injunction banning the union's action because no prestrike vote had been held. However, on the following day the union ordered its members to return to work and talks with management were resumed.

The Post Office made modest concessions and on April 6 what was described by both parties as "an historic agreement" was signed. This 50-page document provided for the unimpeded use of new technology, a more reliable postal service through the better distribution of manpower, and savings estimated at £180 million. The staff were promised no compulsory layoffs, a one-hour reduction in the work week, and a share in the cost savings that could mean additional bonuses of £7 to £14 a week. In other words it seems that the use of the law, with the possibility of substantial fines for the union, was a crucial factor in the union's decision to sign a model agreement that would lead to higher productivity, lower costs, higher pay, and a better postal service for the customer.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

By mid-May 1985 another dispute was threatening, which could have had unpleasant consequences for millions of Londoners and a large number of visitors to the capital. London Regional Transport (LRT) had stated its intention to withdraw the 1984 pay and productivity award of 7.5 percent from 7,000 staff because of their refusal to cooperate with the extension of one-man trains on the underground. Drivers of one-man trains can earn up to £30 more in an average week. Most of the staff belong to the National Union of Railwaymen, and the NUR called a strike for May 2. Following the strike call, LRT was granted an injunction against the union on the grounds that the strike had not been approved in a prestrike vote. The strike went ahead, but it was poorly supported and London Regional Transport was able to operate a reduced service in all areas. On the following day, NUR leaders voted to suspend the strike, publicly admitting that the legal action had deterred some railwaymen from joining the strike. Subsequently, Dr. Tony Ridley, head of LRT, stated that it would not be proceeding against the union and confirmed that the introduction of more one-man trains would not entail compulsory layoffs.

*The Threatened Dispute on the Railways, August 1985*

For several years the British Railways Board, which controls a subsidized monopoly industry, had been trying to persuade the National Union of Railwaymen that guards were unnecessary on most freight and short-distance passenger trains, and some driver-only trains had been introduced. In the summer of 1985, however, the matter came to a head, with the NUR refusing to permit the introduction of more driver-only trains and British Rail insisting the new arrangements had to be accepted. The guards were promised that the new trains would be phased in gradually and that no guards would be laid off involuntarily. In August guards in Scotland and Wales went on strike as a protest against the new manning arrangements, and after refusing to return to work, despite an ultimatum from management, 251 guards were dismissed. The NUR threatened to bring all of the 11,000 guards out on strike, and British Rail made it plain that if that happened it would shut the whole railway system down.

The NUR, however, decided to keep within the law by holding a prestrike vote among the guards; the union, management, and independent observers confidently expected that this would show a large majority favored a strike. In a ballot in which 84 percent of the guards voted, a majority of 52 percent voted against strike action. Three weeks after this surprising result the NUR agreed to take guards off



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

some trains and indicated that the union would be willing to discuss future productivity deals through existing negotiating machinery. The five-year logjam on productivity improvements for Britain's railways had been well and truly broken.

## Effects of the Reformed Law

The pattern of events in the Stockport Messenger dispute, the three disputes of March–May 1985, and the British Rail dispute indicates clearly that the new curbs on trade union immunities can be extremely effective in stopping illegal action and in making trade union officials pay heed to the views of their members. But that is not to suggest that the law will always be effective in preventing disputes. Strikes about pay and conditions of work are still perfectly legitimate where they have been approved in a secret ballot by the employees concerned. Clearly the attitude of ordinary union members is now crucial and their attitude will depend on a number of factors, including the reasonableness of the employer and the state of the economy, especially the levels of inflation and unemployment. The law is an important factor, but it is certainly not the only factor.

The "historic agreement" eventually reached between the Post Office and the Union of Communication Workers in April 1985 was clearly the result of many months, if not years, of careful planning by a management determined to reduce costs (especially labor costs) and to improve services without jeopardizing the position of its existing employees. There is a very important lesson to be learned from this kind of creative negotiation, and unless the lesson is taken to heart there will be no cure for the British disease. It is a lesson that some are slow to learn because, as Artus (1984) has pointed out, labor costs in Britain (as well as France and Germany) are still too high. Recently there has been much talk but not enough action, although a partial remedy for this problem was clearly spelled out more than 20 years ago.

## The Fawley Productivity Agreements

It was in 1964 that the case study by Allan Flanders of management and collective bargaining, *The Fawley Productivity Agreements*, was published. In 1960 the Esso refinery at Fawley was the scene of an unprecedented departure from conventional collective bargaining. The company offered its employees large increases in wage rates in return for their trade unions' consent to specific changes in working practices, including a drastic reduction in overtime, some relaxation of job demarcations, and the redeployment of craftsmen's mates.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The laudable aim was to raise productivity and distribute its benefits without causing an inflationary rise in costs. This aim was largely achieved. Flanders estimated that "the two year increase in productivity was a little less than one half" (that is, 45–50 percent). But the most important part of his classic study was not concerned with the details of the negotiations at Fawley, interesting as they were. It might fairly be claimed that chapter 6, entitled "Lessons of Wider Application," is one of the most significant comments that has been written on the poor state of British labor relations, and the consequent low labor productivity, since the Second World War.

The fundamental problem at Fawley was underemployment, and the corollary of underemployment is, of course, low labor productivity. The chief symptom of underemployment at Fawley was excessive overtime, and the question Flanders posed and answered in his concluding chapter was simply, "Who was responsible for the excessive overtime?" There was no doubt in Flanders's mind about the answer: "The responsibility for reducing it [overtime] rests on the same authority that has been responsible for its growth, namely management" (p. 228). He went on to say that "what makes the [Fawley] experiment stand out in the post-war industrial scene is that, whether or not management found the best ways of changing restrictive practices, *it did not evade its responsibility for making the attempt* and the attempt was not made in vain" (p. 238, emphasis added).

Finally, in section 4 of chapter 6 Flanders considered the "Roots of Management Irresponsibility" and argued persuasively that those roots were to be found in management itself. He said that the slowness of British managers to innovate in the field of labor relations could be explained by the fact that most of them preferred to have as little as possible to do with labor relations and normally delegated them to a personnel manager or department. And he concluded that most senior British managers were distinctly underqualified in terms of education and experience for their job. This view was recently confirmed by Crockett and Elias (1984), who found that 52 percent of a sample of 2,637 British managers had no qualifications at all and only 3 percent were university graduates.

Is there, then, a conflict between Flanders's view of widespread management irresponsibility in British business and Hayek's insistence on the need to reduce or eliminate trade union immunities? Not necessarily. Hayek has argued that radical changes in trade union law are a *necessary* condition for economic recovery in Britain. Nowhere has he argued that reform of the law is a *sufficient* condition for economic recovery, or that it will guarantee, of itself, a British



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

economic miracle. But in the resolution of the Post Office dispute in April 1985 we can see just what could be achieved in the economy as a whole if intelligent, fair-minded, and determined managers were prepared, when the occasion demanded, to use a reformed law to achieve their aims. And in that and other disputes (especially the Stockport Messenger dispute) it can be seen that Hayek was quite right. If the unions had not to some extent been deprived of their coercive powers, through a reduction in their legal immunities, the employers could not have resisted them as they did.

It is, then, reasonable to argue that a substantial improvement in the quality of British managers—especially line managers involved in labor relations—is one of the preconditions for the regeneration of the British economy and at the same time to insist that this must go hand in hand with a reform of trade union and labor law. It may even be that obsolete labor laws have contributed to the poor quality of management by demoralizing existing managers and deterring potential managers of ability from entering what they have seen as an industrial jungle dominated by militant trade unionists. However, it is not the purpose of this paper to examine proposals to improve the quality of British managers, desirable though such an improvement may be. Rather, in the concluding section, the emphasis will be on what now needs to be done by managers and government, following the labor law reforms of 1980–84, to assist the regeneration of the British economy.

## Conclusion

### *Lessons for Managers*

Perhaps the clearest lesson to be drawn from the labor disputes analyzed above is that management firmness, expressed in a willingness to use a reformed law, often helps to produce a satisfactory settlement. In particular Section 10 of the Trade Union Act of 1984, which removes immunity from trade unions that call out their members on strike without a secret prestrike vote, has enabled managers to insist on the right of individual employees to express a personal view about the desirability of strike action. Certainly there is no guarantee that a prestrike vote will produce a vote against strike action. But if managers are willing to use the reformed law against trade unions that flout it, it fully disposes of the argument that a strike has been concocted by a handful of politically motivated agitators.

Nevertheless, it would be extremely foolish to assume that the law is the ultimate answer to labor relations problems in Britain. In their recent study of 23 successful British companies, Goldsmith and Clut-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

terbuck (1984, p. 59) wrote that each of those companies "extracts far more from its employees in terms of personal commitment, dedicated hard work and involvement than the typical British company would think possible," and they went on to show how the companies obtained such devotion from their employees. Until the typical British firm follows the example of these exceptional companies and deliberately pays more attention to achieving a much higher level of employee commitment, there will be no general improvement in the performance of the British economy. The appropriate use of a reformed labor law on its own is not enough.

*Lessons for Government*

Following the failure of the Industrial Relations Act of 1971, the quite illogical argument that the reform of British trade union law was impossible was sometimes put forward and too readily accepted. Recent events have shown the fallacy of this argument. It is, therefore, sad and surprising that the government now seems determined to call a halt to its program of labor law reform instead of capitalizing on its success.

The three acts of 1980, 1982, and 1984 all amend the legislation of the previous Labour government. Thus the statute law is now in an extremely complex, not to say unintelligible, state. There is an urgent need for a consolidating act or acts that would simplify and amend the law. In the 1970s, and especially in the winter of 1978–79, Britain suffered from disputes in several essential monopoly public services. Employees in those services have behaved more responsibly in recent years, but the freedom to strike in these services, following a ballot in favor of such action, might be said to encourage irresponsible behavior. In fact in 1983 the Conservative government promised in their election manifesto (p. 12) to "consult further about the need for industrial relations in specified essential services to be governed by adequate procedure agreements, breach of which would deprive industrial action of immunity." That promise has yet to be kept.

At the same time, Section 1 of the Employment Protection Act of 1975 charges the state Conciliation and Arbitration Service (ACAS) with the duty "of encouraging the extension of collective bargaining and the development and, where necessary, reform of collective bargaining machinery." The British government has given open or tacit support to the extension of collective bargaining since the final report of the Whitley Committee in 1918, but the generally poor state of labor relations and low labor productivity suggest that the policy has failed. Surely it is time for the state to adopt a neutral stance toward collective bargaining and to allow firms to adopt whatever



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

labor relations policies they judge to be in their best interests, while protecting the employee's right to join a trade union or not. A reduction of immunity for unions in essential services and the adoption of a neutral stance by the state toward collective bargaining are just two of the reforms that are still desirable.

When he pointed out the need for radical rethinking about the rule of law in the field of labor Hayek (1960, p. 268) wrote, "The acquisition of privilege by the unions has nowhere been as spectacular as in Britain." The long overdue reforms of 1980–84 have helped to bring the British economy back from the brink of disaster. It would indeed be tragic if the debate about labor law reform were now regarded as closed. Instead the recent reforms should be seen as the basis from which Hayek's (1960, p. 284) ideal—"the strict prevention of all coercion except in the enforcement of general abstract rules equally applicable to all"—can be achieved. If that were to be done, a genuine British economic miracle would surely be probable rather than possible.

## References


Allen, G. C. *The British Disease.* Hobart Paper 67. London: Institute of Economic Affairs, 1976.

Artus, Jacques R. "The Disequilibrium Real Wage Rate Hypothesis: An Empirical Evaluation." *IMF Staff Papers* 31 (June 1984): 249–302.

Brown, A. J. "Inflation and the British Sickness." *Economic Journal* 89 (March 1979): 1–12.

Crockett, Geoffrey, and Elias, Peter. "British Managers: A Study of Their Education, Training, Mobility and Earnings." *British Journal of Industrial Relations* 22 (March 1984): 34–46.

Dicey, A. V. *Lectures on the Relation between Law and Public Opinion in the Nineteenth Century.* 2d ed. London: Macmillan, 1914.

Flanders, Allan. *The Fawley Productivity Agreements.* London: Faber and Faber, 1964.

Goldsmith, Walter, and Clutterbuck, David. *The Winning Streak.* London: Weidenfeld and Nicolson, 1984.

Hayek, F. A. *The Constitution of Liberty.* London: Routledge and Kegan Paul, 1960.

Hayek, F. A. *1980s Unemployment and the Unions.* Hobart Paper 87. London: Institute of Economic Affairs, 1980.

Minford, Patrick. *Unemployment, Cause and Cure.* Oxford: Martin Robertson, 1983.

Pollard, Sidney. *The Wasting of the British Economy.* London: Croom Helm, 1982.

Pratten, C. F. *Labour Productivity Differentials within International Companies.* Cambridge: Cambridge University Press, 1976.

Rowley, Charles K. "Towards a Political Economy of British Labor Law." *University of Chicago Law Review* 51 (Fall 1984): 1135–60.




LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Roy, A. D. "Labour Productivity in 1980: An International Comparison." *National Institute Economic Review* 101 (August 1982): 26–37.
Royal Commission on the Press 1961–1962. *Report*. Cmnd. 1811. London: H.M.S.O., 1962.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Ethnic Utopianism and Market Reality in South Africa

*Nelson Reid and Gary Lowe*

The recent writing on South Africa generally expresses a sense of urgency regarding the necessity to dismantle apartheid and get on with a new regime. The analyses and discussion have often been of the "describe and deplore" variety in which South African social policy is excoriated as the prime example of white racism in the world. Typically, such analyses are based on a presumption of a total linkage of interests between the ruling classes of government and economy in a propertyless, mobile, and willing black labor force. What is rarely acknowledged, however, is that apartheid is a distinctly political creation that has caused an adversarial relationship between the powerful, centralized South African government and the business sector. Consideration of the essential elements of apartheid (more recently termed "separate development") and its compatibility with the usual pattern of capitalist development makes this adversarial relationship understandable.

The apartheid system of social development is conceptually based on a "national" model of group relations in which each "people" of South Africa is presumed to be a "nation" in search of a geographical, cultural, and political base. But the South African economy, still the most robust and diverse on the continent, offers no divisions that would give life or reason to such a political vision of separate nations. Such an extraordinary model of social development as apartheid could only be the product of grand ideology and a powerful political will. That ideology and will are contained in Afrikaner nationalism.

---

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

Nelson Reid is Professor of Sociology and Anthropology and Director of the Social Work Program at North Carolina State University. He was Visiting Professor of Social Welfare at the University of Cape Town for a portion of 1984. Gary Lowe is Lecturer in Social Work at the University of North Carolina at Charlotte and was a Visiting Senior Lecturer at the School of Social Work at the University of Cape Town during 1983–85.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

This paper explores the nature and force of the Afrikaner nationalist movement that has been the source of the utopian vision of apartheid. This vision is the fundamental source of the disjuncture of economic and social reality that separate development requires. While the central question is how to disentangle the South African economy from the politics of apartheid and still leave the economy whole and sound, it is our view that the alternatives facing South Africa are best understood in the context of the country's history of ethnic nationalism. Apartheid is a specific set of social and political policies that reflects a peculiar construction of reality grounded in Afrikaner experience and myth. The successful future of South Africa, we hold, depends not only on the dismantling of separate development policy and the end of the Afrikaner political monopoly, but also on a reliance on free markets and a nonobtrusive government to restore economic stability and racial peace.

## Apartheid/Separate Development

Apartheid has a long history in South Africa. It is based on a central historical obsession with ethnic differences and assumptions about the nature of people as forming communities that possess permanent elements of "culture" that cannot, and should not, be eroded. This view took root long before the 20th century, but the basic elements of the assumptions about culture and community and the basic outline of apartheid are apparent in this century as early as the 1903 Native Affairs Commission report (Fredrickson 1981; Cell 1982). That report, informed in large part by the anthropology of the time, elaborates the necessity for a governmental policy that would preserve native African culture and protect white European culture (Tatz 1962).

Since 1903, the assumptions about culture and people and the perceived necessity to protect European culture, or its specific South African derivatives, have been refined into a social program that reached its height in rhetoric and governmental action during the early period of National Party rule, from 1948 to the mid 1960s, in the governments of Malan, Verwoerd, and Vorster. This period represents the culmination of the Afrikaner ethnic "revolution."

The basic elements of apartheid are those establishing the social, economic, and political context in which blacks may operate. These include the Black Land Acts (1913) specifying "homeland" areas reserved "exclusively" for occupation and tenure by Africans; the Blacks Consolidation Act (1945) creating influx control and pass law requirements for blacks outside homeland areas; the Prohibition of



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Mixed Marriages Act (1949); the Immorality Act (1957) prohibiting sexual relations between whites and blacks; the Population Registration Act (1950) requiring classification into white, coloured, or black (revised to include several subgroups of coloured and African categories); the Promotion of Black Self-Government Act (1959) and the related National States Citizenship Act (1970) establishing all Africans as citizens of a self-governing "national" state; the Group Areas Act (1966) specifying the geographical areas outside the National States in which racial groups may reside, own or rent property, or work; and the Separate Amenities Act (1953) requiring segregation of public accommodations and services.

A set of political controls was added to this scheme of social engineering: the Political Interference Act, which prohibited mixed race political parties or activities; and the Suppression of Communism Act (1954), which allowed the Prime Minister (now State President) to name as a communist anyone who "advocated, advised, defended, or encouraged . . . any of the objects of communism" either actively or by "omission which is calculated" to advance communism.

The result of this grand design has been to reduce blacks to the status of guest workers, complete with non-South African citizenship passports and work and residence permits. This elaborate structure was constructed in order to preserve the fiction of a white European republic and to effect a functional one-party state controlled by Afrikaners via the National Party. Undoubtedly, the differential socioeconomic status of white and black South Africans is not entirely a product of social planning and policy, but these policies have certainly made a substantial contribution toward maintaining the concentration of political and economic power among whites. There have been, however, some recent changes in the policy and administration of the apartheid system.

## Recent Developments

During 1984 and 1985, South Africa erupted with widespread internal disorder. Although the country has had regular periods of political unrest, particularly since the Sharpeville incident in 1960, the last major period of unrest occurred in 1976 primarily in Soweto. This recent unrest, though repeating many of the aspects of previous incidents, is distinct in two critical ways. First, the protests, school boycotts, and other related acts have been sustained over many months and have spread throughout the country. Second, the participant groups in Soweto represented a broad spectrum of South Africa's disenfranchised, not simply the black African majority population



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

groups. This was most evident in the Western Cape and Cape Town's coloured areas, in that many of the protestors were "middle class" and invested more materially in South Africa than had their black African counterparts. This would seem to bring into question the government's ability to build alliances with "elite" coloured and Asian groups that have benefited from economic growth in South Africa over the past decade. The government responded to the unrest by declaring a limited State of Emergency in July 1985 and extending it to Cape Town and the Western Cape in November. The State of Emergency remained in effect, overtly at least, until March 1986.

The unrest in Soweto followed four years of steady, if not always profound, reform efforts by the current National Party government headed by State President P. W. Botha. The government's reforms include legalization of black labor unions in major industries; constitutional changes creating three legislative houses, one white, one coloured, one Asian (or Indian); the removal of the Political Interference Act; partial desegregation of restaurants, theaters, sporting events, and railways; the repeal of the legal prohibition of mixed marriages; changes in administration of Group Areas legislation allowing some business presence of blacks in central urban areas; and, most recently, a promise of unified South African citizenship for all and suspension of the much despised pass laws. Many of these reforms are either cosmetic or a matter of bringing law into conformity with reality, but a few of these reforms promise to have far-reaching effects.

The constitutional changes creating houses of representation for coloured and Asian voters have been widely criticized for excluding black African representatives. Indeed, the current unrest in the country is directly related to the passage and implementation of the new constitution. But the "new dispensation," as the government likes to call it, incorporates governmental opposition into the national legislative system in content and scale quite out of character with previous National party governments. However limited a reform, it does move the South African government into a more modern governmental context in which the State President must act, in part, as a broker among competing political interests in order to be effective.

The governmental concept that has held sway in the National Party since it gained power in 1948 is one that emphasizes government as the extension and administrator of "truth," endowed with a sacred cultural mission (Adam 1971). Such a government is "good" and its opposition is "evil," so tolerance of political diversity is predictably low. The current governmental structure, however, seems to admit the possibility of legitimate differences of interest that can be incor-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

porated into an institutional framework. The problem of black political representation will certainly be much easier for the government to deal with in the context of a pluralist-modern view of government as broker as opposed to the previously dominant view of government by "revelation."

The legalization of black labor unions in the mining industries has given rise to legitimizing black political and social interests expressed in the context of labor. Because South Africa is a nation in which there is considerable state ownership and participation in critical industries, the legalization of labor unions for blacks is, in fact, an avenue for political participation. The major black labor organizations are strongly anti-apartheid and anti-National Party and have recently joined forces in the potent Congress of South African Trade Unions (COSATU).

The April 1986 announced intent to abolish the pass laws also has potentially far-reaching consequences. The "influx control" measures are designed to regulate the movement of black labor to urban areas and prevent black workers from becoming permanent residents of areas officially designated as white. Such measures are often spoken of by the government as being necessary for orderly movement, but there is no evidence that the laws have substantially impeded geographic mobility to urban areas. They have certainly been the most intrusive aspect of apartheid for blacks, with arrests for pass law violations reaching into the hundreds of thousands by the early 1980s. The pass laws represent a central element in the apartheid system and most visibly carry the idea of blacks as guest workers in their own nation. The importance of their elimination would be second only to the abolishment of the Group Areas Act.

## Allocation of Resources Among Population Groups

*The Race Relations Survey 1984*, prepared annually by the research staff of the South African Institute of Race Relations (SAIRR), reported that South Africa's population (including all 10 of its homelands) was 32.6 million in mid 1984, with black Africans making up 72.8 percent of the total. Table 1 provides a breakdown of total population groups and shows the urban-rural distribution, and Table 2 shows demographic, welfare-related indicators of the South African population. Table 3 profiles the average monthly earnings and the household incomes for the different race groups for 1980–84, and Table 4 illustrates changes in public employment income by racial group from 1970 to 1981.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 1*

THE URBAN RURAL DISTRIBUTION OF SOUTH AFRICAN POPULATION, 1980

| | Urban | | Rural | | Total | |
|---|---|---|---|---|---|---|
| Asians | 743,820 | 90.6[a] | 77,500 | 9.4[a] | 821,320 | 100[a] |
| Africans | 6,809,660 | 32.6 | 14,094,100 | 67.4 | 20,903,760 | 100 |
| Coloureds | 2,002,300 | 76.6 | 610,480 | 23.4 | 2,612,780 | 100 |
| Whites | 4,002,000 | 88.4 | 526,100 | 11.6 | 4,528,100 | 100 |
| Total | 13,557,780 | 47.0 | 15,308,180 | 53.0 | 28,865,960 | 100 |

[a]Percent
SOURCE: Marais (1985, p. 23).

These data indicate a pattern of resource allocation that is replicated in many sectors crucial to South Africa's national development. Although in general inequalities between racial groups have declined during the last 20 years, the continuing major differences are a well-known dimension of South Africa. Some of these inequalities are directly a product of governmental action, others arguably are not; but to construct social policies and programs and to allocate the associated benefits along racial boundaries is in evident violation of the political philosophy and practice established in western societies. The usual explanation has been to describe the South African ruling white oligarchy as being concerned only with the ascendancy of whites and cruelly unconcerned with the fate of others. But "racism" as it is commonly understood is a painfully simple and inadequate concept in the South African context, and to understand the development of the peculiar vision of apartheid it is necessary to understand the context of ethnic (Afrikaner) nationalism in which it developed.

## Ethnic Utopianism and Afrikaner Nationalism

Arguments can be presented that the social development impetus and the supporting theory and method that have been so prevalent in the decolonial half of this century contain rather explicit utopian assumptions. Though South Africa's modern social development contains elements from the existent and mainstream developmental paradigms, it seems to owe more to 19th-century utopianism than to 20th-century developmental notions.

For many people the affective connotation of "utopia" will not fit well with the overall image that is assumed to be associated with South Africa. Traditionally, and in ordinary usage, "utopia" connotes



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 2*

DEMOGRAPHIC INDICATORS OF THE SOUTH AFRICAN POPULATION

| | Asians | Whites | Coloureds | Blacks |
|---|---|---|---|---|
| Birth Rate,[a] 1982 | 25.1 | 16.6 | 30.0 | 38.5 |
| Death Rate,[b] 1982 | 5.9 | 8.3 | 9.9 | 11.0 |
| Percentage Annual Growth Rate, 1970–80 | 2.7 | 1.8 | 2.4 | 3.0 |
| Total Fertility Rate,[c] 1980 | 2.7 | 2.0 | 3.3 | 5.3 |
| Life Expectancy at Birth, 1980 | 64.8[g] | 70.3[g] | 58.5[g] | 57.5[g] |
| Infantile Mortality,[d] 1982 | 17.6 | 13.6 | 58.4 | 90.0[g] |
| Dependency Ration, 1980: | | | | |
| Children (<15)[e] | 61.8 | 42.9 | 65.2 | 70.9 |
| Aged (>65)[f] | 4.3 | 12.1 | 5.1 | 5.4 |

[a]Number of births in a particular year per 1,000 of the population in the middle of that year.

[b]Number of deaths in a particular year per 1,000 of the population in the middle of that year.

[c]Denotes an index indicating the total number of children that a woman will give birth to if the fertility pattern for the particular calendar year or period does not change in the course of time.

[d]Calculated by expressing the number of deaths occurring among infants under the age of one year during a particular calendar year in relation to the number of live births during that year. The rate is expressed per 1,000 live births.

[e]Population 0–14/population 15–64 × 100.

[f]Population 65+/population 15–64 × 100.

[g]Estimates (mainly by the Institute for Sociological and Demographic Research of the HSRC).

SOURCE: Marais (1985, p. 23).

a positive ideal: "Utopia as [the] ideal . . . the perfect society, has long exerted a powerful influence upon the thinking, feeling, and action of human beings" (Richter 1971, p. 11). But apartheid was very much a positive ideal pursued chiefly by Afrikaner and academic social engineers of South Africa who held with a firm belief in the greater good of separate development for all "groups." If utopia involves "the willed transformation, in which a new kind of life has been achieved by human effort" (Williams 1980, p. 196) and the "utopian mentality [is] a state of mind that is incongruous with the state of reality within which it occurs" (Mannheim 1936, p. 192), then South African development is well within the utopian tradition.

From the perspective of the ruling National Party, the South African nation and the apartheid policy represent a "willed transforma-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 3*

AVERAGE MONTHLY EARNINGS AND HOUSEHOLD INCOMES, SELECTED YEARS

| | Average Monthly Earnings[a] | | | |
|---|---|---|---|---|
| | African | Asian | Coloured | White |
| 1979 | R156 | R278 | R210 | R665 |
| 1980 | R189 | R336 | R254 | R767 |
| 1981 | R228 | R412 | R309 | R936 |
| 1982 | R272 | R513 | R365 | R1,093 |
| 1983 | R310 | R584 | R417 | R1,210 |

| | Household Incomes[b] | | | |
|---|---|---|---|---|
| | African | Asian | Coloured | White |
| 1980 | R136 | R505 | R344 | R912 |
| 1981 | R174 | R591 | R417 | R1,134 |
| 1982 | R204 | R819 | R548 | R1,380 |
| 1983 | R220 | R934 | R558 | R1,630 |
| 1984 | R273 | R1,072 | R624 | R1,834 |

[a]The National Manpower Commission gave these figures regarding average monthly earnings of workers in all sectors of the economy excluding agricultural and domestic service.

[b]The figures for the claimed monthly average household income (in money terms) for the different race groups were acquired from Market Research Africa.

SOURCE: Cooper et al. (1985, p. 241).

tion" that has been explicitly implemented since 1948. These policies and programs rest on a foundation and a rationale anchored in an overt historical, social, and religious mythology developed prior to 1948.[1] From the perspective of outside observers, as well as a significant portion of the South African population, these developmental policies and programs and their proponents surely possess a "state of mind" that is wed to a vision "incongruous with the state of reality."

Mannheim (1936, p. 191), among others, has observed the disruptive character of utopian ideas:

> [T]hose orientations transcending reality [are] Utopian . . . [and] when they pass over into conduct, tend to shatter, either partially or wholly, the order of things prevailing at the time.

[1]See Moodie (1975), Adam and Gilliomee (1979), van den Berghe (1967a, 1967b), Hexman (1981), and Adam (1971).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*TABLE 4*

WAGES AND SALARIES OF PUBLIC EMPLOYEES BY RACE GROUP, SELECTED YEARS

| | Whites | | | Coloureds | | | Asians | | | Blacks | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number Employed | Average Earnings R.p.a. | Wage Index | Number Employed | Average Earnings R.p.a. | Wage Index | Number Employed | Average Earnings R.p.a. | Wage Index | Number Employed | Average Earnings R.p.a. | Wage Index | Benefits Accrued to Lower Income Group R'000 |
| 1970 | 99,313 | 3,262 | 100,00 | 30,745 | 1,350 | 100,00 | 7,304 | 1,745 | 100,00 | 135,465 | 536 | 100,00 | 108,779 |
| 1975 | 106,768 | 5,304 | 162,60 | 42,227 | 2,335 | 172,96 | 8,706 | 3,885 | 222,69 | 114,686 | 1,280 | 238,81 | 238,023 |
| 1981 | 129,858 | 9,240 | 283,26 | 60,089 | 4,178 | 309,48 | 13,011 | 8,437 | 483,50 | 132,064 | 2,842 | 530,22 | 626,424 |

SOURCE: *South African Statistics 1982* (Pretoria: Central Statistical Services).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Goodwin and Taylor (1982, p. 138) made a similar observation:

> Utopias are not to be defined simply as idealistic "wish-images"; they are wish-images which, *when put into practice,* shatter the existing order and work instead to establish an alternative system in which they believe they [and possibly everyone else as well], can achieve complete happiness and fulfillment.

During this century, the utopian goal has become a secularized end attainable by conscious human efforts rather than an other-worldly ideal brought about through divine intervention. Utopian political movements have very much expressed the "will to power" that has been such an apparent feature of 20th-century social development (Berger 1976; Johnson 1983). This is not to say, however, that religious rationales do not provide important support for concrete secular policies and actions, as the South African case reveals quite clearly.

The specific content of the utopian vision that has dominated in South Africa is manifest in the social, political, and cultural intentions of Afrikaner nationalism. That national movement is based on a concept of the historic cultural mission of the Afrikaner people. The essence of the controlling idea is expressed in the following quotations:

> History so decided that of all countries of Africa and Asia colonized from Europe there was to be one, and only one, that cradled a new nation with a new language and a distinct culture, allied to but at the same time different from its countries of origin [Pienaar 1960, p. 3].

> White nationalism and Black nationalism must be diverted into confluent channels. There must be a White South Africa and Black South Africa politically divided but peacefully and cooperatively coexistent [Pienaar 1960, p. 12].

> Apartheid comprises a whole multiplicity of phenomena. It comprises the political sphere; it is necessary in the social sphere; it is aimed at in church matters; it is relevant to every sphere of life [Davenport 1977, p. 270].

In 1948, the National Party won a narrow electoral victory, placing it in control of the South African government. The ascendancy to political power signified by this electoral victory was the culmination of a steady process stretching back into the 19th century when the Afrikaners in the Cape Colony conducted their "Great Trek" to establish the Orange Free State and the Transvaal. Although the bitter Boer War was lost to the British, Afrikaner nationalism was never effectively suppressed; by the 1930s, it had reached great social influence with its emphasis on Afrikaners as "poor whites" with a distinct and valuable culture.

The years following World War II and the Nationalist political victory were marked by a rapid consolidation of political, bureau-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

cratic, and economic power by Afrikaners through their National Party. Prior to 1948, numerous legislated antecedents had defined the nature of inter-group and ethnic relations in South Africa. Primary among these were the Natives Act of 1923 which stipulated residential segregation between blacks and whites; influx control measures for urban areas; and the explicit assertion that Africans were to be "visitors" to white man's territory. During the mid 1920s and into the 1930s, a Civilized Labour Policy developed that gave whites preference, particularly in urban areas and in mining and industrial employment. Much of this policy was implemented as a result of labor union pressure, which was itself tied to the larger Afrikaner nationalist movement. In 1936, the Cape Province voting franchise for blacks was removed. All of these actions were relatively uncoordinated, but they served as the underpinnings for what would ultimately emerge as the set of social development policies and programs subsumed under the lable of apartheid. As W. H. Hutt (1975, p. 57) noted, "Race discriminations in S. Africa had their *origins* in history . . . [but] their *perpetuation* has been due to the quite natural use of State power and trade union collusion to preserve the status quo." Since 1948, the apparent interest of policy has been to perpetuate the inferior status of non-whites in an "almost sadistic affront to the dignity and self-respect of the non-white races."

Between 1910 and 1948, a process of mythology-building was initiated that unified the indigenous frontier Afrikaner colonizers. This development of Afrikaner nationalism and its attendant social vision promoted the post-1948 social development in philosophy and program. Afrikaner nationalism as a force has evoked extensive debate and speculation concerning its foundations and its precise content. There is relative consensus that religion, specifically Dutch reform doctrine, was critical; it combined with inter-group conflicts and a significant colonial experience to become a particular political and social ideology. Economic factors, especially the discovery of gold and diamond deposits and the emergence of a labor movement during the 19th century, also contributed to the emergence of Afrikaner nationalism. The complex interplay of these elements combined during the 20th century to create an Afrikaner sacred history to explain the past and a doctrine of civil religion that would provide cultural energy for the present (DuToit 1983).

After the Afrikaner elective victory in 1948, the basic elements of the "national" model of ethnic relations set into place before World War II were extended and refined through the set of interlocking laws and regulations discussed earlier. The peak of political support for this nationalist social development model was reached during the



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tenure of Prime Minister H. F. Verwoerd from 1958 to 1966. Verwoerd gave apartheid a philosophical and conceptual structure, and his forceful and charismatic personality greatly assisted in the political struggle to carry these ideas forward. Verwoerd, former head of the Department of Sociology at Stellenbosch University, was able to breathe life into the myths and aspirations of Afrikaner nationalism. In his roles as Minister for Native Affairs and then Prime Minister, he achieved the logical social development program implications of that myth. It is notable that academics dominated the "age of social engineers," as T. R. H. Davenport refers to it.

Thompson (1985, p. 1) has provided a useful delineation of myth that is relevant to South Africa and the role myth plays in the political arena:

> [P]olitical myth [is] a tale told about the past to legitimize or discredit a regime; and . . . political mythodology [is] a cluster of such myths that reinforce one another and jointly constitute the historical element in the ideology of the regime or its rival.

Afrikaner nationalism is mythological in form and is used in support of utopian ends. Labeling Afrikaner nationalism as mythological is not a matter of judgment but of description; it recognizes the status of this central aspect of the ideology that has informed this century's social development in South Africa. Thompson (1985, p. 8) made another useful observation regarding the nationalism mythology and the reach for utopia it has engendered: "Myths originate in specific circumstances as a product of specific interest, and they change within the changing interests of successive generations and successive regimes."

The myth of Afrikaner nationalism contains three major elements (Moodie 1975): sacred history, civil theology, and civil ritual. Sacred history consists of history given meaning by God and, therefore, it carries a clear sense of purpose and mission for the Afrikaner Volk. This history has unified Afrikanerdom and has provided the rationale for the social and economic policies promulgated by the Volk once temporal power was acquired. This sacred history is replete with communal suffering and martyrs, and most of these events and people have served to provide historical evidence for the notion of Afrikaners as a chosen and blessed people.

Afrikaner theology is anchored to an Old Testament God and to Calvinist-related doctrine. As Moodie (1975, p. 11) pointed out, "The Lord of Afrikanerdom is sovereign and intensely active, present at every turning point in the affairs and nations of men." This civil theology was functionally powerful as it served to "explain" the



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

communal sufferings; in this way it linked the past to the present and offered sure promise for the future. The "civil" portion of the theology was the Republican ideal, the rule of the people represented in the two Afrikaner republics defeated and absorbed as a result of the Boer War. This was "the suffering," which would presumably pass and lead to the reestablishment of a new "nation" if the Afrikaner Volk remained cohesive. As Moodie (1975, p. 17) concluded, "[E]verything which emphasized Afrikaner uniqueness—their language, their Calvinist faith, their customs and conventions, their very dress—took on sacred significance."

Just as the United States and other social and political entities have developed civil rituals as occasions to assert a communal bond, Afrikaners developed these rituals from both sacred history and civil religion. These rituals served to assert and solidify the Afrikaner's claim to cultural and communal uniqueness. For Afrikanerdom, several monuments serve as the focus for the civil ritual and represent its sacred history: the Blood River *Laager*, the *Vrouemonument* (National Women's Monument) near Bloemfontein in Orange Free State, and the *Voortrekker* Monument overlooking Pretoria. These physical entities stand like the monuments in Washington, D.C., as reminders of a not so distant past and its meaning for the present; each one commemorates an Afrikaner victory over forces that would undo them.

The effective establishment of Afrikaner history, civil religion, and ritual took less than half a century. All of these elements were developed, actively promoted, and combined so that they "united Afrikaners in their sense of unique identity and destiny, inspired the faithful, converted the skeptical, and even reminded them of *their sacred separation from English and black African*" (Moodie 1975, p. 21).

So it was that this foundation was firmly in place when the Nationalists gained political control of South Africa in 1948. On this complex and fervently held world view and supporting belief system the Afrikaner nationalists set out to establish their utopian vision. It was from this perspective that social development drew its energy and quite logically took the form of apartheid/separate development with its emphasis on "peoples" as "nations."

As Banton (1983, p. 232) has observed, Afrikaners see themselves "as a Nasie (nation) or Volk, rather than a race . . . and the arguments they use to defend their policies depend on assumptions about the distinctiveness of peoples and the essential qualities of nationhood, and not upon typological or selectionist theories of racial inequality." The Afrikaner nationalist movement sought, and obtained, political legitimacy and power through the National party. Similar to political



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

vehicles of such movements elsewhere, the party and the people it represented were presented as one: "die party is die volk and die volk is die party." There is precious little room in such a view for the sort of tolerance of political variation more typical in western democracies. The governments of the National Party have never seemed comfortable with the idea of government as broker between competing political interests; they have preferred the notion of government by "revelation" and "truth." The Botha government acknowledges more overtly than any since pre-war Union Party Smuts that it must bargain, but it has not been able to do so gracefully.

South African history during this century is best understood as a product of ethnic revolution. As Adam and Giliomee (1979) argued, it is a revolution that has successfully captured and enhanced a powerful central government, it has informed that government's policies with an ideology of Afrikaner nationalism, and it has pursued policies and programs that establish Afrikaner culture and people as recognizably superior in important symbolic and material ways.

## Ethnic Nationalism, Government, and the Free Market

If South African history is viewed in terms of ethnic revolution with powerful elements of a history of domination and a distinctive religious and ideological system, then South African history is perhaps not so unique as is commonly believed. There is evidence of a somewhat similar mixture of government, ethnicity, and cultural idealism in other places in the world, including elsewhere in Africa (Meredith 1985). Cyprus, Northern Ireland, Lebanon, Pakiston, Sri Lanka, and India are all nations that have experienced ethnic conflict; in these countries, control over government has been an element in institutionalizing or often exacerbating the conflict. In Africa, it is the rare nation in which ethnic, ideological, and religious boundaries do not have great political meaning. If South Africa is a "state without a nation," unable to establish a central and common identity that conforms to its political boundaries, then the same must be said for many African nations that have found genuine nationhood to be elusive.

The phenomena of disharmony in racially and ethnically pluralist societies are so widespread that certain generalizations are possible. There are numerous theories of race relations, many of them variations on the themes of capitalist development and class conflict; but some place special emphasis on the distinctive roles of economy and politics in producing harmonious or disharmonious relations among



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

people. Michael Banton (1983), for example, applied a variation of social exchange theory to argue that individual competition dissolves group boundaries and group competition hardens group boundaries. Because individual competition occurs in economic markets and group competition occurs primarily in the political realm, often requiring the government to be an agent of group interest, the argument implies that the role of government in race relations is generally negative.

In a review of the process of social "boundary maintenance" in Uganda, Northern Ireland, Britain, Malaysia, the United States, and South Africa, Banton (1983) developed his "rational choice theory" of race relations. He argued strongly for the position that racial politics, even when perceived to be favorable to "minority" groups, tends to harden racial identity and worsen human relationships. Alvin Rabushka (1974, p. 69), basing his argument on the economic theory of market relationships and public goods theories, came to a similar conclusion: "Racial tensions and conflicts are kept at a minimum under conditions of voluntary exchange in free markets." Banton's and Rabushka's arguments would seem to question whether or not any overtly racial policy that is the product of group political competition would not undermine community relationships.

Many researchers have found evidence of the central role of governmental policy in institutionalizing, conflictual ethnic group relations.[2] This would seem to support Banton's and Rabushka's conclusion that group competition is fostered in politics and that politics in multi-racial societies becomes all too often a vehicle for racially defined interests. Following this logic, for racial harmony of governments in multi-group societies to exist, there must be a decentralized structure that does not lend itself to ethnic "take-over," a scale of government that would not reward racial politics, and reliance on an economy largely free of governmental control.

The South African economy is sufficient to produce a per capita income in excess of $3,000 (more than three times Zimbabwe). The Republic draws over 1.5 million black workers from its southern African neighbors. In addition, the government has created and funded education, health, and welfare programs that in structure and purpose look much like those in the developed West. The distribution of benefits under such programs is skewed unacceptably; but if the franchise is broadened, the probability is that the distribution of governmental benefits will become more equal (placing an even

[2]See Morrison and Stevenson (1972), Becker (1957), Degler (1971), Fogel and Engerman (1974), and van der Berghe (1967a, 1967b).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

higher tax burden on the already rather highly taxed middle and upper income groups). One cannot resist the observation that the South African government has the opportunity to move toward a repudiation of apartheid and social engineering and to embrace a decentralized, less powerful, and less important government that will oversee a free economy and a free people. To do so, however, it must give up its ethnic nationalist dreams, be content with the Afrikaner's place as one of many, and adopt the position that what Afrikaners foisted upon South Africa no subsequent ethnic group shall be in a position to repeat.

## Reform and the Business Sector

Apartheid is a political creation, a utopianism made impossible by economic forces that inevitably run counter to the establishment of geographic and political communities based on ethnic identity. One would expect, therefore, that the business community, despite its traditional apolitical character, would be less than supportive of apartheid. In the South African context, there are rather extensive and well-organized business groups that have little interest in centralized governmental authority or apartheid. As is evident by the recent delegation of South African businessmen to Lusaka to visit with Oliver Tambo and other leaders of the African National Congress (ANC) as well as the frequent pleas for reform that emanate from the national South African Chamber of Commerce, businessmen are interested in political stability, a non-politically polluted investment climate, a non-politicized labor force, and increasing disposable income for all South Africans. The "middle"—those who seek a non-racial, less-intrusive government—may not be so apparent in South African politics, but it is readily visible in South African society as any visit to Cape Town, Johannesburg, or Durban will attest.

In recent years, business firms, particularly international firms, have had a high profile in reform efforts. Perhaps the major example to date has been the support American firms (especially Mobil) have given to the Sullivan Principles. Lesser examples are General Motors' promise to provide legal support for its employees in Port Elizabeth who are prosecuted for violating beach segregation laws and Coca Cola's $25 million trust fund for black education. In addition, American businesses in South Africa have contributed more than $80 million to community work. This may appear to be moral protection money, but virtually all business publications, including the influential *Financial Mail* and *Business Day*, support the dismantling of apartheid and the creation of a broadly representative government.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Gavin Relly, chairman of the Anglo-American mining conglomerate, recently said that apartheid is an "albatross" that has "demeaned the white man as much as it has degraded the black man" (*Raleigh* [N.C.] *News and Observer*, June 1, 1986). In June 1986, the American Chamber of Commerce in South Africa, under the leadership of such companies as IBM, Coca-Cola, Citibank, and Union Carbide, ran full-page newspaper ads calling for sweeping reforms and pronouncing apartheid "totally contrary to the idea of free enterprise" (*Raleigh News and Observer*, June 3, 1986). A recent survey of over 900 business leaders by Project Free Enterprise, based at the University of Pretoria's School of Business Leadership, found that only 17 percent supported the Botha government's process of negotiated change and 93 percent endorsed a political structure inclusive of all groups in the central governent (*Charlotte* [N.C.] *Observer*, June 9, 1986). Because of the current state of things, many American firms have threatened to pull out of South Africa, and quite a few have already done so. Some of the concern is due to disinvestment pressure by educational institutions and public employee groups, and some is due to the worsening economic climate; but much also seems based on the view that the social engineering of apartheid is just not compatible with modern economic development.

The business sector wants reform, but it does not want another Mozambique or Angola or Zimbabwe. It wants what is difficult to find elsewhere on the continent: a stable, non-racial, non-intrusive government for a prosperous market-oriented multi-ethnic society. The business-oriented do not seem to embrace any precise social vision that is to be brought about by the government. What is needed, according to the business community, is for the natural forces of commerce and culture to work free of governmental direction. But such a viewpoint does not lend itself well to South Africa's charged political climate.

Conor Cruise O'Brien recently wrote that "African Socialism has no success stories to tell." If socialism is in essence the confusion of society and state and the presumption of the power and responsibility of the state to fashion a society, then South Africa is surely in both the socialist and African tradition. It is troubling, however, that one of the principal devices used to influence South Africa may undermine the very sector of the society that has the greatest appreciation for the governmental and social structure that would most likely produce social harmony.

Disinvestment, widely advocated and increasingly implemented, is a strategy that chooses as its initial target private business and corporate interests. But what is its purpose? Is it the intention of



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

disinvestment to prod the business sector to higher levels of political activism? Is it based on the view that the economic and governmental sectors are some sort of interlocking directorate? Or is it simply founded on the Marxist-derived notion that governmental policy in South Africa has reflected business class interests in a mobile, willing, powerless, and large supply of black labor? If disinvestment is effective it will cause general economic harm in South Africa, slowing economic growth and retarding labor opportunities. If disinvestment is successful, it will reduce the presence and importance of international business in South Africa and punish the economic sector generally. Literature and logic would seem to hold that racial harmony is not enhanced under such circumstances. Apartheid is not a creation of the economic sector; it is a politically imposed and quite artificial arrangement that surely would not flow from the workings of an unhindered economy.

## Conclusions and Implications

In his "Afrikaner Nationalist Perspective," John Seiler (1979, p. 3) wrote:

> Whatever change comes about in South Africa—short of a prolonged civil war augmented by external invasion—will be defined by Afrikaner nationalist perceptions, shaped by Afrikaner national interests, and processed by institutions monopolized by Afrikaner nationalists.

From the perspective of several years later, his assertion seems somewhat less certain. Granted that anti-South African governmental interests get more rhetorical than material support from either the West or the East, and granted that South African society is still very much intact; still, both the pace and the character of change seem quite different from that of the late 1970s. Which direction or directions South Africa will take is beyond the power and knowledge of most to predict, but certain lessons might be drawn from the country's social and political history.

If the Afrikaner ethnic "revolution" model is accepted as an explanation of the development of apartheid and the specific centralized/utopian character of the South African government and if the rational choice theory of race relations is accepted, the implications for South Africa's future would seem to be rather clear: The political structure must allow expression of various and conflicting interests, but it must be structured in a way that will not allow similar ethnic or other narrowly ideological movements to gain such authority. In essence, a less powerful, more decentralized governmental sructure without



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

an overarching social vision or ideology (and without the power to implement it) combined with freedom of labor, contract, and investment would have the greatest promise of establishing a government most able to fulfill its capability of achieving a high standard of welfare for its people, including the establishment of political freedom, civil rights, and long-term stability as a more just and nonracialist society.

## References


Adam, Heribert. *Modernizing Racial Domination.* Berkeley: University of California Press, 1971.

Adam, Heribert, and Giliomee, Hermann. *The Rise and Crisis of Afrikaner Power.* Cape Town: David Philip, 1979.

"America and South Africa." *The Economist,* 30 March 1985.

Banton, Michael. *Racial and Ethnic Competition.* New York: Cambridge University Press, 1983.

Becker, Gary. *The Economics of Discrimination.* Chicago: University of Chicago Press, 1957.

Berger, Peter. *Pyramids of Sacrifice: Political Ethics and Social Change.* Garden City, N.Y.: Anchor Books, 1976.

van den Berghe, Pierre L. *Race and Racism: A Comparative Perspective.* New York: John Wiley, 1967a.

van den Berghe, Pierre L. *South Africa: A Study in Conflict.* Berkeley: University of California Press, 1967b.

Cell, John W. *The Highest Stage of White Supremacy: The Origins of Segregation in South Africa and the American South.* New York: Cambridge University Press, 1982.

Cooper, Carole et al. *Race Relations Survey, 1984.* Johannesburg: South African Institute of Race Relations, 1985.

Davenport, T. R. H. *South Africa: A Modern History.* 2d ed. Toronto: University of Toronto Press, 1977.

Degler, Carl N. *Neither Black nor White: Slavery and Race Relations in Brazil and the United States.* New York: Macmillan, 1971.

DuToit, Andre. "No Chosen People: The Myth of the Calvinist Origins of Afrikaner Nationalism and Racial Ideology." *American Historical Review* 88 (1983): 920–52.

Fogel, Robert William, and Engerman, Stanley L. *Time on the Cross: The Economics of American Negro Slavery.* Boston: Little, Brown, 1974.

Fredrickson, George M. *White Supremacy: A Comparative Study in American History.* London: Oxford University Press, 1981.

Goodwin, Barbara, and Taylor, Keith. *The Politics of Utopia: A Study in Theory and Practice.* London: Hutchinson, 1982.

Hexham, Irving. *The Irony of Apartheid: The Struggle for National Independence of Afrikaner Calvinism Against British Imperialism.* New York: The Edwin Mellen Press, 1981.




LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hutt, William H. "South Africa's Salvation in Classical Liberalism." In *Individual Freedom: Selected Works of William H. Hutt*. Edited by S. Pejovich and D. Klingman. Westport, Conn: Greenwood Press, 1975.

Johnson, Paul. *Modern Times: The World from the Twenties to the Eighties*. New York: Harper and Row, 1983.

Mannheim, Karl. *Ideology and Utopia*. New York: Harvest/HBJ, 1936.

Marais, Dr. H. C. *The South African Society: Realities and Future Prospects—HSRC Investigation into Intergroup Relations*. Pretoria: Human Science Research Council, 1985.

Meredith, Martin. *The First Dance of Freedom: Black Africa in the Post-War Era*. New York: Harper and Row, 1985.

Moodie, Dunbar. *The Rise of Afrikanerdom: Power, Apartheid, and the Afrikaner Civil Religion*. Berkeley: University of California Press, 1975.

Morrison, D. G., and Stevenson, H. M. "Integration and Instability: Patterns of African Development." *American Political Science Review* 3 (September 1972): 902–27.

Pienaar, S. "Safeguarding the Nations of South Africa." In S. Pienaar and A. Sampson, *South Africa: Two Views of Separate Development*. London: Oxford University Press, 1960.

Rabushka, Alvin. *A Theory of Racial Harmony*. Columbia: Institute of International Studies, University of South Carolina Press, 1974.

Richter, Peyton E., ed. *Utopias: Social Ideals and Communal Experience*. Boston: Holbrook Press, 1971.

Seiler, John. "The Afrikaner Nationalist Perspective." In *South Africa Into the 1980's*, pp. 3–19. Edited by Richard Bissell and Chester Crocker. Boulder, Colorado: Westview Press, 1979.

Tatz, C. M. *Shadow and Substance in South Africa*. Cape Town: University of Cape Town Press, 1962.

Thompson, Leonard. *The Political Mythology of Apartheid*. New Haven, Conn.: Yale, 1985.

Williams, Raymond. *Problems in Materialism and Culture*. Verso, 1980.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Establishing Property Rights in Outer Space

*Joel D. Scheraga*

> One way or another, the world seems headed toward a legal order to allocate the increasingly scarce and valuable geosynchronous orbits. . . . Only with a wider sense of shared commitment and less national insecurity will the world be able to enjoy the full benefits of space.
>
> —Daniel Deudney

## Introduction

Space is a resource. With the advent and proliferation of space travel, particular locations in outer space are rapidly becoming scarce resources. The demand for satellite systems has grown rapidly during the past decade. Such growth in demand follows the development of new satellite launch systems and reductions in the costs of these systems. With the commercialization of launch services, satellite technology is rapidly becoming less costly to obtain. NASA's Space Shuttle, the European Space Agency's Ariane rocket, and NASA's Atlas Centaur and Delta rocket systems (both soon to be privately controlled) are now competing for public and private customers.[1]

There has concurrently been an increased awareness of the benefits to be gained from an unmanned presence in outer space. Com-

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

The author is Visiting Assistant Professor of Economics at Princeton University and Assistant Professor of Economics at Rutgers University. Thanks are due James W. Head, David Schap, Joseph Seneca, and Howard R. Swearer for helpful comments. The valuable insights of Arieh Eilan, former Israeli Ambassador at Large to the United Nations, also are acknowledged. Earlier versions of this paper were presented at Swarthmore College and at the session on property rights at the 1985 Meetings of the Eastern Economic Association. The comments offered at these sessions proved useful in revising the paper for publication.

[1]NASA has selected Transpace Carriers Inc. of Greenbelt, Maryland, to be the private marketer of launches on the Delta rocket. NASA is also negotiating with General Dynamics to become the marketer of the Atlas-Centaur launch system.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

munications satellites have revolutionized worldwide telecommunications. Weather satellites provide near-instantaneous updates on weather conditions. Reconnaissance satellites provide continuous monitoring of military activities in foreign countries. Teleconferencing services via satellite and direct delivery of television programming to residential areas are becoming commonplace.

The economic benefits from satellite technology are large. But although the demand for satellite systems is rapidly increasing, the orbital paths into which the satellites can be placed are limited. These paths represent a scarce resource because they are desirable but limited. Scarcity implies the need for resource allocation among unlimited human wants—the fundamental problem of economics.

The issue addressed in this paper is the allocation of the limited and desirable physical locations in outer space for geosynchronous satellites. How will these scarce resources be allocated? Can they be efficiently allocated? If peaceful and efficient coexistence in space is to prevail, then the creation and enforcement of property rights is inevitable.

## The Problem: Geosynchronous Satellites

A geosynchronous satellite is one that is placed in a west-to-east orbit over the earth. The satellite is placed at an approximate altitude of 22,300 miles, where its period of revolution around the earth is 24 hours. The orbital movement of the satellite is synchronized with the earth's rotation so that the satellite appears from particular points on earth to remain stationary. Continuous transmission of information can be attained between those points. Electronic channels, known as transponders, are built into the satellite and receive, amplify, and retransmit microwave signals sent from earth.

### *Technical Aspects of the Scarcity Problem*

The physical locations in space where geosynchronous satellites can be placed are limited. To fully appreciate why this is true, it is necessary to understand the technical characteristics of satellite systems—specifically, the limitation of the number of radio frequencies for transmissions and the required spacing of satellites in geosynchronous orbit.

Satellite systems consist of several primary components: the satellite in geosynchronous orbit, the on-board transmission system, and the earth-based antenna and transmission station. Satellite broadcasting relies on the use of the radio-frequency spectrum. Different types of transmissions have different bandwidth requirements within



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the spectrum. For example, voice and data transmissions require only small bandwidths within the spectrum. Television transmissions have much more extensive bandwidth requirements. The number of bandwidths, however, is limited. It is possible for many satellites to use the same frequencies for their transmissions, but satellites must nevertheless be spaced sufficiently far apart from other satellites using the same frequency in order to guarantee that interference in transmissions will not occur. The spacing requirement is sensitive to the available power supply on board the satellite and the size of the earth-based antennae. Current technical considerations require that each satellite be placed approximately 2–3 degrees away from any other satellite using the same transmission frequencies. This means that a maximum of 180 orbital "slots" would be available to satellites using the same frequencies and occupying the same 360-degree orbital path.

The usefulness of available orbital slots is further limited because not all slots are equally suitable for particular regions. In particular, the orbital arc over the equator north or south of the region to be served is the most valuable. These orbital arcs generally fall between the longitudes of the major continents. Consider, for example, the arcs of the orbital paths that are of greatest interest to the United States, Canada, Mexico, and Latin America. The orbital arc of interest to the United States lies between 60 and 135 degrees west longitude because satellites in this area can serve the entire continental United States. Satellites in the western 20 degrees of the arc can also be seen from Hawaii and Alaska. But this same 75-degree arc is optimal for satellite communications in Canada and Mexico. The eastern portion of the arc also contains the optimal locations for Latin America. But with spacing of 2 degrees, only 38 communications satellites using the same transmission frequency can occupy this 75-degree arc.

Congestion is inevitable if the demand for stationary satellites continues to grow. Some of the orbital arcs over parts of the equator south of Europe and North America are already crowded. INTELSAT, the international telecommunications satellite organization, projects that slots in some orbital arcs will be filled by the early 1990s (Deudney 1982).

The technical aspects of satellite transmissions, therefore, lead to two significant questions about the limited availability of orbital slots for geosynchronous satellites: (1) Are the existing frequency allocations that are in the desirable portions of the radio-frequency spectrum adequate, given the expected demand for satellite systems? (2) Even if the number of frequency allocations were adequate, would the number of orbital slots be sufficient?



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*Offsetting Effects*

Although such technical constraints as frequency availability and spacing requirements contribute to the allocation problem, several other technical factors help mitigate the problem. First, the 2-degree spacing requirement applies only to satellites that use the same radio frequencies. In practice, satellites that use different frequencies can be placed adjacent to one another. In this case, the spacing requirement is reduced to the minimum distance deemed necessary to ensure that the satellites do not collide. (The FCC currently requires spacing of only one-tenth of a degree to reduce the probability of a collision.)

Second, using existing technology, it is possible to have satellites focus their transmissions into small geographic areas. As radio technology improves and the area of focus is reduced, the possibility of signals from adjacent satellites overlapping will be reduced, even if the transmissions are on the same frequencies. As the diameter of radio beams is reduced (for example, by using laser beams), the required spacing between satellites will also be reduced, thus increasing the capacity of any geosynchronous orbit.

Finally, a single satellite communications system cannot transmit and receive signals on the same frequency at the same time because interference will occur. Each satellite must therefore use at least two frequencies to receive and transmit signals to earth. But the capacity of the geosynchronous orbit can be doubled by having a second communications satellite in the same orbit use the same frequencies but for transmissions in the opposite directions (Martin 1978).

## An Economic Description of the Problem

*Congestion in Space: A Failure to Assign Property Rights*

Private property does not yet exist in outer space. In fact, steps have been taken on the international level to prevent its establishment. In 1967 the U.N. Committee on the Peaceful Uses of Outer Space, which is responsible for legislating all matters dealing with space, drafted an Outer Space Treaty.[2] Ratified by 107 member nations, the treaty provides the current framework for space law. The central principle underlying the treaty is that outer space is not subject to appropriation by any one country; all nations have equal rights and access to the resources of space. Outer space is the "province of all humankind" (Cowen 1985). The treaty's failure to establish property rights is critical. Congestion in geosynchronous orbits will worsen

[2]The title of the draft is "Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies."



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

in the absence of property rights. The price of an orbital slot is zero, so that entry into synchronous orbit is free.

The argument for establishing property rights in outer space is an application of what McCloskey (1985, p. 330) has called Adam Smith's generalization: If transactions costs are low, the assignment and voluntary exchange of rights to scarce resources will result in an efficient allocation. Conversely, the failure to assign property rights to the scarce resources will inevitably lead to an inefficient use of the resources.[3] The problem is a common one in economics. Consider, for example, the overhunting of the buffalo on the Great Plains. The opportunity cost of hunting the buffalo, in terms of yet-to-be-born buffalo, was zero. They were overhunted and killed almost to the point of extinction because no one owned them. The few remaining buffalo survived only because laws that established property rights to the remaining buffalo and their unborn offspring finally protected them (McCloskey 1985, pp. 330–31).

As applied to outer space, Smith's generalization implies that an efficient use of scarce orbital slots will result once property rights are assigned unambiguously to a particular country (or coalition of countries) and free exchange is permitted so that the country can sell the property rights for whatever the market will offer. A common counterargument is that the nations of the world, operating in their own self-interests, will conserve the orbital slots even in the absence of well-defined property rights. But this argument is mistaken: if the price of an orbital slot is zero and the orbital paths are not owned by anyone, the opportunity cost to any one nation of occupying these locations is lower than if property rights were assigned. Orbital paths for geosynchronous satellites will be overused by individual countries and congestion problems will worsen. External costs to firms and nations that may want subsequently to occupy these orbits will not be fully taken into account. The problem is one of ownership.

*The Inevitability of Property Rights in Outer Space*

Economic theory suggests that property rights will be created when it is in someone's self-interest to do so. Demsetz (1967) has argued that the emergence of property rights takes place in response to the desires of individuals (or governments) for adjustment to new cost-benefit possibilities. According to Demsetz (p. 350): "Changes in knowledge result in changes in production functions, market values, and aspirations. New techniques, new ways of doing the same things, and doing new things—all invoke harmful and beneficial effects to

[3]This principle is often referred to as Coase's theorem in the literature (see Coase 1960).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

which society has not been accustomed." The advent of satellite technology and communication was a major technological change. New markets were created—markets to which old property rights were no longer applicable—and the relative prices of the different techniques changed dramatically.

The emergence of property rights as a result of the new technology has been a gradual process. During the early days of space exploration, the United States and the Soviet Union were the only nations that possessed the technology to place satellites into geosynchronous orbit and relatively few were launched. Since orbital locations for synchronous satellites were limited and only a few satellites were launched, it did not pay for anyone to be concerned with the allocation of orbital slots. Property rights were not defined. The development of new satellite technology and more sophisticated and cost-competitive rocket delivery systems has increased the rate at which orbital slots are being filled. The value of the synchronous orbital paths, therefore, has increased.[4] Property rights to the orbital slots will be established as it becomes worthwhile for interested nations to be concerned with the establishment and allocation of such rights. A set of economic and social relations specifying the rights of each country to use the orbital slots will be created in order to allocate these scarce resources in specific and predictable ways.

The issue is scarcity, and the particular locations in space that must be occupied by geosynchronous satellites are scarce. They must somehow be allocated. The market will ensure that property rights to the scarce orbital slots are defined and enforced.

*Private Property Eliminates Congestion*

Congestion of orbital slots is the result of an absence of ownership. Consider the case of a country that has decided to invest in a satellite communications system rather than, say, a ground-based microwave transmission system. When that country's satellite is placed in geosynchronous orbit, it adds to the congestion problem and increases the possibility of transmission interference or collision with another satellite. Although the external effect on each individual satellite in the orbit is small, the total effect on all satellites is large. The country launching the new satellite, however, does not consider the total external effect on all satellites; that is, it does not consider the social cost of one more satellite being placed in orbit. It only considers the

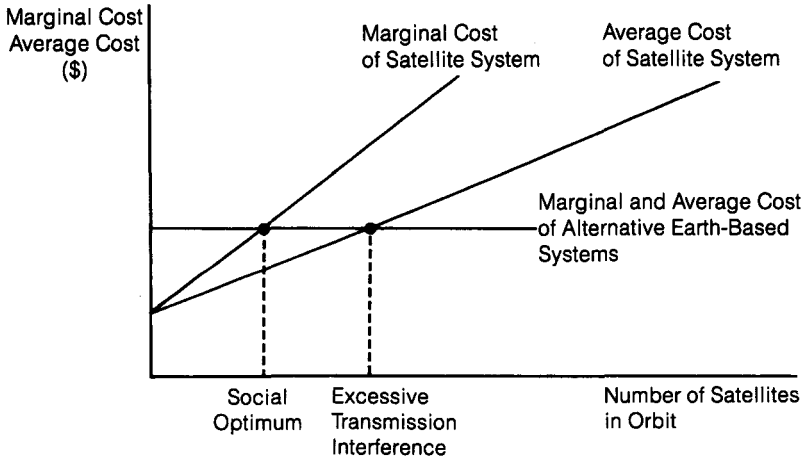
[4]A market analysis by Batelle Laboratories of Columbus, Ohio, predicts that the different launch systems will be competing for 400 to 700 commercial payloads from now through 1998.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

average cost (or cost per satellite launched) it faces—that is, the private cost of the satellite system. Each individual country acting alone, in its own self-interest, will not make socially correct decisions when the orbital slots are not owned by anyone. This misallocation due to the lack of well-defined property rights is illustrated in Figure 1.

*FIGURE 1*
THE SOCIAL OPTIMUM




Under the status quo, orbital slots are not owned by anyone and the price of a slot is zero.[5] A country that is contemplating placing a communications satellite into a geosynchronous orbit will only consider the average cost of the system. It will choose the satellite system over the alternative ground-based systems until the average cost of placing satellites in orbit is equal to the marginal cost of the alternative systems.[6] Satellites will be placed in orbit up to the point where the average cost of the satellite system (the private cost) is equal to the marginal cost of alternative, uncongested systems. Although this is an optimal private decision, it is inefficient for society. The social optimum is obtained when investment by different countries in satellite systems is allocated so that the marginal cost of the earth-based systems is just equal to the marginal cost of the satellite system.

[5]It is interesting to speculate whether country A could, with legal impunity, knock country B's satellite out of an orbital path that A wishes to occupy. It would appear (as is later discussed) that there is a set of de facto property rights established on the basis of a first-come, first-served allocation scheme.

[6]The marginal cost of an alternative system is equal to the change in total cost that occurs when one more ground-based system is built.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The social optimum, of course, is unknown ex ante; rather, it will tend to emerge once private property rights to orbital slots are assigned and enforced. By assigning property rights, a market is established in which the rights to the orbital slots may be bought and sold. Selfish maximization of the profit from property rights will lead to a socially efficient outcome. The negative externalities will be eliminated. The owner of a right to an orbital slot will charge a positive price for the slot that maximizes his net revenue. This price will be the one that induces countries to recognize the costs they impose on others by adding to the congestion of satellites.

If all orbital paths are owned and transactions costs are low, an efficient outcome will prevail. It does not matter which country initially obtains the right to a particular orbit. If exchange is costless, the right will eventually be owned by the country that values it the most. As Cheung (1970, p. 64) noted:

> Competition for and transferability of the ownership right in the market place thus perform . . . main functions for contracting. . . . [C]ompetition conglomerates knowledge from all potential owners—the knowledge of alternative contractual arrangements and uses of the resource; and transferability of property rights ensures [via flexible relative prices] that the most valuable will be utilized.

If the market is allowed to operate, then an efficient (although not necessarily equitable) outcome will prevail.

A classic example of the importance of private property for achieving a socially efficient use of resources is the distribution of property rights in the United States to radio and television frequencies (Coase 1959). In the 1920s there were no restrictions on who could broadcast on any frequency. Chaos ensued. Consequently, the courts adopted a first-come, first-served method of allocating frequencies. The first user of a frequency had ownership claims to it. In general, this type of allocation scheme did not result in frequencies being owned by those users that valued them the most. Nevertheless, if the rights to the frequencies could have been traded, an efficient outcome would have prevailed. A redistribution of wealth would have occurred in favor of the initial owners of the frequencies, but the final allocation would have been efficient. The U.S. government, however, chose to allocate the frequencies according to its own criteria.

## Alternative Systems of Allocation

### *The Historical Record*

The historical record of attempts to resolve the congestion problem is not encouraging. The international community recognizes that a



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

problem exists, but agreement on its resolution has been difficult because of differences in national objectives and in the present endowments of technology and wealth. The United States and the Soviet Union both recognize the need for international agreement on an allocation scheme. They also recognize, however, that such an agreement might force them to share satellite technology and give up orbital slots to Third World nations. Orbital slots have thus far essentially been allocated on a first-come, first-served basis.

The United Nations' International Telecommunications Union (ITU) currently attempts to allocate the slots among countries. But the ITU's judgments are not enforceable, and not all countries recognize its jurisdiction. Many Third World nations view the rapid appropriation of orbital slots as an inequitable and inefficient distribution of resources. Developing countries oppose the allocation process because they believe many of the slots will be filled by the time they can use the satellite technology. The United States supports the allocation scheme and argues that the orbital slots really are not scarce; that technological innovation will improve transmission systems and increase the capacity of the orbits. However, a congestion problem already exists and an efficient system of property rights needs to be put in place now.

Several countries have opposed the U.N. Outer Space Treaty and the notion that outer space is not subject to national appropriation. In 1976, seven equatorial countries (most directly affected by the most desirable orbital slots) signed the Bogota Declaration, staking a claim on the orbits directly over their countries. They argued that because geosynchronous satellites are in a stationary position over their countries, the orbits are an extension of their territorial space. Although this declaration directly conflicts with the Outer Space Treaty and is not recognized by most other countries, the question of where national airspace becomes international space has not been resolved.

More recently, India has proposed a licensing system for the orbital slots. The proposal calls for each country to be awarded a certain number of slots. These licenses could then be bought and sold. The advantage of this system is that it would permit the use of the orbits now and would also provide minimum property rights to developing countries.

Economists will recognize in these conflicting attitudes and proposals the fundamental conflict between efficiency and equity. Systems of allocation exist that will inevitably lead to efficient outcomes. There are few disagreements over the efficiency or inefficiency of



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

these systems. The major source of contention is over the equity of alternative allocation schemes (Scheraga 1986a, 1986b).

*Types of Ownership and Enforcement Costs*

Various types of ownership are possible, including communal ownership, private ownership, and state ownership (Demsetz 1967). It has already been demonstrated that communal ownership of geosynchronous orbital paths is not feasible. Individual users of satellites do not have an incentive to limit the rate at which they occupy the orbital paths. Although, in theory, it is conceivable that the "international community" could jointly own these rights, the costs of reaching such an agreement would be very high. The costs of negotiation would be high because it would be difficult for countries to reach a mutually satisfactory agreement. This situation is aggravated by the unequal distribution of wealth, technology, and satellite launch capabilities. Even if an agreement could be reached, the costs of monitoring and policing it would be high.

The property rights, therefore, must be either privately or state-owned. Since the market for satellite technology and launch systems is international in scope, agreements on the definition and ownership of the property rights must be made at the international level. Governments must negotiate the allocation of orbital locations and agree on different allotments. The ITU has, in fact, attempted to administer geosynchronous orbits through international meetings and negotiations, but the union does not recognize the competitive market mechanism as a means of allocating the slots. Under the current nonprice allocation scheme, then, the ITU holds the market price to zero and allows political competition to determine who gets the available slots.

It is likely that if an individual government recognizes private ownership, the allotment will somehow be placed in the hands of private entrepreneurs. (Within the United States, the Federal Communications Commission has the responsibility of allocating orbital slots to commercial users.) Alternatively, if private ownership is not recognized, then the rights will be state-owned. It is essential for the efficient distribution of these property rights that the cost of exchanging the rights be low relative to gains from trade. In this case (neglecting the effects of the distribution of wealth on the valuation of orbital slots), it does not matter who initially obtains the orbital property rights. After negotiation and exchange, the rights will end up in the hands of those countries and firms that value them the most.

The result of private ownership of the property rights will be to internalize many of the external costs associated with communal



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ownership. Each private owner will have an incentive to use his resource efficiently and to take into account the value of the property right.[7]

*Alternative Allocation Schemes*

The current method of allocation does not achieve an efficient or an equitable outcome. Under this system, the orbital slots are not owned, and there is no incentive for countries—especially those with large endowments of technology and wealth—to conserve the slots by restricting their use. The nonexistence of property rights has led to a congestion problem. If the slots were owned, the resource would be better used. Given that the existing system is not efficient, what alternative methods of allocation would be preferable?

*A Bidding System.* Adam Smith's generalization states that an efficient outcome will prevail if all orbital slots are owned and transactions costs are not prohibitive. This general principle provides a compelling argument for a simple bidding system—that is, a free-market economy in which mutually beneficial exchanges can occur.

The difficulty with an auction market is that the wealthiest countries would obtain most of the property rights and therefore would affect the outcomes of future trades between countries. A deal between any two countries must yield for one country a higher level of utility than before the trade and leave the second country no worse off, if the deal is to be mutually acceptable and is to take place voluntarily. But the extent of mutually beneficial trades will depend on the initial endowments of the two countries. It is therefore important to question the equity of the initial distribution of orbital slots.[8] (Of course, if an auction system did exist and the price of slots was relatively low, the equity issue would no longer be a serious concern.)

[7]The right of ownership consists of the right to use a scarce orbital slot, to sell all of the rights to the slot, or to transfer some of the rights through rental of satellite services. The property rights to transponders on individual, privately owned communications satellites are already being sold. For example, AT&T leases transponders on its Comstar D-4 satellite to other telecommunications corporations. Hughes Communications Services Inc., a subsidiary of Hughes Aircraft Co., intends to lease to the Navy, for military communications, a satellite launched from the space shuttle Discovery.

[8]The Indian proposal for a licensing system hints at a solution to the equity problem. A more sophisticated organization or international agreement would have to be set up to award, with equity considerations in mind, a certain number of orbital slots to each country. Once the licenses were distributed, countries could trade and an efficient outcome would prevail. The problem with this system is the difficulty of getting all countries to agree on the initial distribution of slots.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*The International Court of Justice.* As suggested by the Bogota Declaration, questions of international law enter into the allocation problem. As the judicial arm of the United Nations, the International Court of Justice, or World Court, has the authority to rule on the legality of claims to the orbital slots. But only 44 of the U.N.'s 159 member states recognize the World Court's jurisdiction, and a majority of the 44 nations attached reservations to their acceptance. Further, the United States and the United Kingdom are the only permanent members of the U.N. Security Council that have accepted the Court's jurisdiction. The World Court has no practical way of compelling any nation to comply with its rulings. The Court is also highly politicized, so that an objective ruling about slot allocation could not be expected.

The seven equatorial countries that signed the Bogota Declaration did so in the belief that their claim of ownership was founded in international law. If the rulings of the World Court were truly binding, the equality of such claims would depend on these rulings. It is doubtful, however, that such rulings could ever be enforced. Despite the World Court's inability to enforce rulings on the initial distribution of orbital slots, Adam Smith's generalization still applies: if exchange is costless, countries will still bid for and trade orbital slots until they end up in the hands of those countries that value them the most.

*Quotas and Licensing.* Another approach to the allocation problem is to introduce quotas on the number of satellites each country can place into orbit. By international agreement, each country would be permitted a fixed number of satellites. The trading of rights to the slots would not be permitted. Quotas on the output of satellites by any country would be similar to quotas imposed on commercial airlines.

Quotas are licenses that have value. They bestow on some lucky few the privilege of earning rents from the orbital slots. A quota system would only coincidentally lead to an efficient and equitable outcome, however. A quota is a special case of a restriction on entry into a market, and restrictions usually lead to inefficient outcomes.

*A Congestion Tax.* If it is too costly to enforce property rights in outer space—that is, if transactions costs are prohibitive—what should be done? Two extreme solutions are to ban entry into geosynchronous orbit (which is unenforceable) or to permit unlimited entry into orbit (essentially the status quo). But these two extremes are inferior to a congestion tax, a solution most economists also prefer over quotas. A tax on each geosynchronous satellite could be imposed that



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

brings the marginal private cost of the system into equality with the marginal social costs. The size of the tax should reflect the amount of money that other countries would be willing to pay to prevent further congestion of the geosynchronous orbit.[9] The congestion tax, therefore, becomes a substitute for establishing private property rights. If a country chooses to orbit a satellite despite the tax, then it is apparently worth more to the country orbiting the satellite than it is worth to other countries facing the increased problem of congestion.

A tax on congestion would appear to result in an optimal number of satellites being placed in orbit at any point in time. But there is a flaw in this argument. One is tempted to conclude that the proper place to impose the tax is on the country orbiting the satellite. But the country "causing" the congestion is not always the proper place to impose the tax. The very existence of other countries interested in occupying the orbital slots can also be viewed as a cause of the congestion problem. The obvious cause of the congestion problem is not always the best place to impose the tax burden.

When the costs of making deals between countries are low and trade is possible, it does not matter where the tax burden is placed. Adam Smith's generalization guarantees that an efficient outcome will still prevail. But if the costs of making deals are high, then Coase's theorem that it *does* matter where the liability is imposed supersedes Smith's generalization.[10] Placing the tax burden on the country orbiting the satellite may lead to an inefficient outcome. Consider, for example, that total world income would be higher if it were cheaper for the countries facing the congestion to find other means of communication than it would be for the country orbiting the satellite to pay the tax. It would be inefficient to impose a tax on the country orbiting the satellite when total world income is maximized by other countries using alternative methods of communication. Regardless of where a congestion tax is imposed, it is still a substitute for private property, and if transactions costs are low, a system of congestion taxation can lead to an efficient and equitable outcome.

[9]It is difficult to imagine how this amount would be determined. For this reason, many economists believe that the imposition of taxes to deal with the externalities caused by congestion inevitably leads away from the social optimum. No public authority, no matter how well intentioned, can compete with a market system in evaluating the effects of externalities (Davis and Whinston 1962). Negative externalities due to congestion are dealt with, in a free market, by the existing owners paying potential owners of orbital slots not to orbit geosynchronous satellites.

[10]In fact, the property rights literature emphasizes the idea that externalities are associated with transactions costs, including the costs of exchanging and enforcing property rights (Demsetz 1964, 1966).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

## Conclusion

The market system applies to scarce resources in outer space as well as to resources on earth. Scarce orbital slots for geosynchronous satellites can be efficiently allocated if property rights are assigned and exchange is permitted. The assignment of property rights, if abided by and done in a spirit of cooperation, need not be feared. Clearly defined property rights underlie well-functioning markets, which are socially beneficial.

This paper has dealt with the issues of why the establishment of property rights to geosynchronous orbital slots is inevitable and how these property rights can be created. The focus has been on economic efficiency. The analysis of alternative allocation schemes suggests that competitive markets can be applied usefully to the problem of allocating scarce orbital slots for geosynchronous satellites.

Clearly defined and enforced property rights provide sufficient incentives to preserve scarce resources. Scarcity in the absence of ownership leads to congestion, inefficiency, and a misuse of resources. Failure to assign property rights leads to socially undesirable outcomes. Policy issues that are concerned with equity, as distinct from efficiency, involve interpersonal comparisons of utility and social welfare judgments on the part of bureaucratic authorities. Further research will be required to investigate the equity implications of establishing property rights in outer space, given the current international distribution of wealth and technology.

## References

Cheung, Steven N. S. "The Structure of a Contract and the Theory of a Non-Exclusive Resource." *Journal of Law and Economics* 13 (April 1970): 49–70.

Coase, Ronald H. "The Federal Communications Commission." *Journal of Law and Economics* 2 (October 1959): 1–40.

Coase, Ronald H. "The Problem of Social Cost." *Journal of Law and Economics* 3 (October 1960): 1–44.

Cowen, Robert C. "Shackling Lunar Developers." *Technology Review* 88, no. 2 (February/March 1985): 6, 77.

Davis, Otto A., and Whinston, Andrew B. "Externalities, Welfare, and the Theory of Games." *Journal of Political Economy* 70 (1962): 241–62.

Demsetz, Harold. "The Exchange and Enforcement of Property Rights." *Journal of Law and Economics* 7 (October 1964): 11–26.

Demsetz, Harold. "Some Aspects of Property Rights." *Journal of Law and Economics* 9 (October 1966): 61–70.

Demsetz, Harold. "Towards a Theory of Property Rights." *American Economic Review* 57 (May 1967): 347–59.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Deudney, Daniel. *Space: The High Frontier in Perspective.* Worldwatch Paper 50. Washington, D.C.: Worldwatch Institute, August 1982.

Martin, James. *Communications Satellite Systems.* Englewood Cliffs, N.J.: Prentice Hall, 1978.

McCloskey, Donald N. *The Applied Theory of Price.* 2d ed. New York: Macmillan, 1985.

Scheraga, Joel D. "Curbing Pollution in Outer Space." *Technology Review* 89, no. 1 (January 1986a): 8–9.

Scheraga, Joel D. "Homesteading and the Creation of Property Rights in Outer Space." In *Space Colonization: Technology and the Liberal Arts,* pp. 114–20. Edited by Charles H. Holbrow, Allan M. Russell, and Gordon F. Sutton. New York: American Institute of Physics, 1986b.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# The Public Domain and Nineteenth Century Transfer Policy

*Gary M. Anderson and Dolores T. Martin*

## Introduction

During the 19th century, about 871.2 million acres of lands held in common by the federal government were transferred to private individuals, businesses, and state governments. The privatization of public lands began in 1796, when the public domain amounted to approximately 233 million acres. Between 1796 and 1923, private parties purchased nearly 279.3 million acres of public land. Even though the total amount of federal land holdings grew enormously, the government sold off more acreage than had been included in its original holdings. Most of these sales occurred before 1862; from 1800 until the beginning of the Civil War, proceeds from the sale of public lands constituted a major source of revenue for the federal government, accounting for 48 percent of net receipts in 1836.

Denationalization of the public domain was not limited to privatization via sales. By 1923, nearly 592 million additional acres of the public domain were transferred to individuals, railroad companies, and state governments through grants; the overwhelming proportion of these transfers occurred after 1862. What had begun as a major revenue-generating device for the U.S. Treasury had evolved into an immense transfer program. A system of allocating scarce resources via auction had been transformed into one of political allocation based on rent seeking, with large sections of the public domain being transferred to politically favored groups.

---

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

Gary M. Anderson is Associate Professor of Economics at California State University–Northridge and Dolores T. Martin is Professor of Economics at the University of Nebraska–Lincoln. The authors wish to thank Larry Hunter, Dwight Lee, Roger Riefler, Gordon Tullock, and Vernon Smith for helpful comments on earlier drafts of this paper.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Little attention has been focused on explaining the dramatic shift in federal land policy during the middle of the 19th century. Granting that privatization per se was a great success during the 19th century, why was privatization via market sales supplanted by privatization via political allocation? An extensive body of literature on public sector decision making and output/pricing policies asserts that the behavior of public sector managers is analogous to that of private sector entrepreneurs, who have the same incentive to restrict competition and extract monopoly profits (see Stigler 1971; Peltzman 1976; McCormick and Tollison 1981). The primary difference between the two kinds of decision makers is in the institutional arrangement of the public sector, which forces managers to maximize revenue in order to gain financial advancement and the perquisites of office. The behavior of governmental policymakers with regard to the public domain during the first half of the 19th century seems consistent with this description. As the century progressed, however, politicians and bureaucrats systematically eroded their own revenue base by awarding enormous land grants to private parties.

Is 19th century federal land policy a paradox for the political self-interested, revenue-maximizing model of governmental behavior? Our intention is to examine the historical record to find answers to this question. We will argue that the system of land grants is consistent with the modern economic model of government as a revenue-maximizing "firm" that supplies income transfers to interest groups. Properly viewed, the land grant programs were an unusual form of spending that permitted politicians and bureaucrats to expand transfers during a period when a rapid expansion of *cash* subsidies was not technically or politically feasible. The land grant system is therefore consistent with the modern economic theory of government.

## Origin and Privatization of the Public Domain, 1789–1862

Between 1780 and 1802, the United States government acquired claims to 233 million acres of land. Great Britain ceded to the U.S. government its claim to land south of the Great Lakes as a condition of the 1783 Treaty of Paris, and the states conveyed their western land claims to the federal government. The amount of land held by the federal government continued to grow during the early 1800s, with 523.4 million acres added by the Louisiana Purchase in 1803. By 1860, 685.3 million acres of territory had been added to the public domain.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The nation's founders intended the federal government to use the public domain to generate revenue from proceeds of land sales to the private sector. One of the first actions taken by Congress during Washington's administration was to pass the Public Debt Act of August 4, 1790, which declared that the proceeds from the sale of the public lands "are hereby appropriated toward sinking and discharging the [national] debts" (Gates 1984, p. 37). Administrative delays precluded the organization of actual sales until 1796, when a land law was passed that established an administrative system for public auctions in which tracts of 640 acres were to be sold at a minimum price of two dollars.

During the first year, only 49,000 acres were sold for $112,135, a very disappointing figure (Rohrbough 1968, pp. 19–21).[1] Over the next four years, nearly 1.3 million acres were sold for approximately $1 million. To increase sales, Congress revised the act on May 10, 1800, establishing the Land Office to administer the auction sales, retaining the two dollar per acre minimum but reducing the minimum tract to 340 acres, and extending credit to the purchaser. Sales rapidly increased after 1800 in terms of both volume and proceeds; between May 1800 and June 1820, over 13.6 million acres were sold for $22.6 million, producing average proceeds of about $1.1 million per annum.

The credit system was ended by the Homestead Act of 1820, which reduced the minimum price per acre to $1.25 and the minimum purchase per tract to 80 acres. Following this legislative change, both acreage and revenues greatly increased. The average annual acreage sold from 1820 to 1862 was about 3.3 million acres compared to an average of only 649,983 acres sold per year from 1800 to 1820. The average receipts after 1820 were nearly $3.6 million per year, a 273 percent increase over the average annual revenues of the preceding period (see Donaldson 1884, pp. 196–208, and Hibbard 1924, pp. 100, 103, 106).

After 1820, receipts from land sales became a major component of federal revenues. During 1836, for example, receipts from land sales exceeded 48 percent of total federal revenues. From 1820 to 1860,

[1]Congress faced competition from other sellers, many of whom offered land at significantly lower prices. As of 1812, individuals held some 30 million acres beyond the Alleghenies; Massachusetts, Pennsylvania, and Georgia offered millions of acres for sale. During the 1790s, New York sold 4 million acres at 8 cents per acre and 6 million acres at 1.5 cents an acre. Between 1785 and 1821, Massachusetts sold about 5 million acres at an average price of 20.5 cents per acre (Lebergott 1984, p. 78).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

receipts from land sales averaged 10.8 percent of total federal receipts per annum.[2]

From the program's beginnings in 1796 until 1862, privatization of the public lands via sales to the private sector scored several major successes. By 1862, acreage equaling about 67 percent of the public domain in 1802 had been sold, and land sale receipts provided a significant, although fluctuating, fraction of total federal revenues.

## The Land Grant System

For every acre of land sold between 1796 and 1923, over two acres were transferred to states, firms and individuals in the form of grants. Almost all of this transfer activity occurred after 1862. Whereas the annual average acreage sold between 1842 and 1862 had been about 3.3 million acres, the average between 1863 and 1923 fell to about 2 million (Hibbard 1924, p. 103).

The federal government disposed of almost 592 million acres of public land using several different types of transfers. The most imporant transfer program, involving approximately 213.9 million acres, was the granting of homesteads. Under the Homestead Act of 1862, 160 acres were made available to any adult who could demonstrate occupation of the land in question for at least five years.

The second largest grants were awarded to the railroads. Between 1850 and 1923, approximately 129 million acres were granted to states and corporations for railroad purposes; nearly 37.8 million acres went to states who immediately allocated them to railroad companies, and 91.2 million acres were given directly to railroad corporations (Gates 1968, p. 384). Such grants were originally justified as being a necessary means for ensuring that rail lines would reach the underpopulated West; it had been claimed that without free land such expansion would be economically impossible. By 1862, however, increasingly liberal grants to railroad corporations (directly and indirectly) were being rationalized on the grounds that they represented a necessary public subsidy of all capital and construction costs to these corporations. Millions of acres were granted to railroads with the understanding that those corporations would sell the land at a later time. Hibbard (1924, pp. 261–66) reports that the average price "of substantially all railroad land sold prior to 1881 was $4.76 per acre"; the Chicago, St. Paul, Minneapolis and Omaha Railroad sold 1.5 million acres of land for $5.23 per acre between 1882 and 1907; nine Iowa railroads sold 3.7 million acres prior to

[2]These revenues were pledged by Congress for payments on the national debt, which was reduced from about $83 million in 1800 to a low of $38 thousand in 1835.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

1893 at an average price of $5.50 per acre; and the Illinois Central sold 1.2 million acres for approximately $11.70 per acre.

The third largest grants of public land were to the states for educational purposes, about 90 percent of which occurred after 1860. Between 1803 and 1910, 29 states received 73.2 million acres for common schools. It might seem odd that Ohio received only 724,266 acres while sparsely populated Montana received 5.2 million acres for elementary schools, but many states resold large portions of their grants to private parties to generate revenue (Hibbard 1924, pp. 322–23). The 1862 act allotted another 11.1 million acres to states for agricultural and mechanic arts colleges. Again, the states sold most of this land, eventually obtaining nearly $17.6 million in receipts. Altogether, states were granted almost 97 million acres for educational purposes, the bulk of the land transfers occurring after 1860.

The next largest grants in terms of acreage were the 68.2 million acres in bounty lands that were transferred to veterans of the Revolutionary War, the War of 1812, and the Mexican War. This form of transfer came closest to using the public lands to defray federal expenses, because it amounted to a payment-in-kind for military services rendered. Most recipients of the land grants did not develop the land but resold their claims, which were granted in the form of scrip. The scrip-holder, whether a veteran or someone else, was entitled to a nonrestricted tract of the assigned size from within the public domain.[3]

The final large category of land grants was 63.9 million acres of so-called swamp grants to states. In theory, the idea was to cede to the states swampland and marshland that would otherwise be unsalable so that they could develop the land by draining the swamps and so forth. But most of this land granted was potentially valuable property, a result of the land selection system, which was carried out by agents *appointed by the states themselves*. Hibbard (1924, p. 278) estimates that it was frequently the case that 75 percent of the land claimed was not swampy or even subject to serious flooding.[4] States later

[3]Most veterans were uninterested in pioneering the West and subsequently sold their land warrants. At the same time that the General Land Office was charging a minimum of $1.25 per acre, most of the warrants sold for 25 to 70 cents per acre (Gates 1968, pp. 263–64). Both the federal government and veterans would have been better off if the government had managed its land sales in an efficient, revenue-maximizing manner and had paid veterans cash bonuses higher than the amounts they could expect by selling land warrants individually.

[4]Hibbard (1924) mentions one case in Champaign County, Illinois, in which the state agent listed 22,000 acres as swamp for purposes of a grant that the land office representative found had over 11,000 acres of dry land.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

disposed of most of the swampland transfers by granting the land to railroads and by selling it to raise general revenue (Hibbard 1924, pp. 284–88).

The other categories of land grants disposed of much smaller amounts of land, although they still accounted for several million acres. Grants of 4.6 million acres were allotted for the construction of canals, 3.3 million acres for wagon roads, and 2.2 million acres for miscellaneous river improvements. The bulk of these grants were undertaken before 1850, although most of the wagon road grants were made for military road construction following the Civil War (Hibbard 1924, pp. 236–39). The Timber Culture Act of 1873 transferred about 9.9 million acres to individuals.

Most of the land transferred as grants was carefully selected by the grantees, not the grantor. Homesteaders were basically unrestricted in their choice of tracts, as were railroad corporations who established the routes they wanted to develop and then appealed to Congress or state legislatures for the necessary grant. Bounty land "vouchers" allowed holders to select their own parcels of land, and swamp grants were manipulated by the states. If we assume rational behavior on the part of the beneficiaries of land grants, it follows that they made a conscientious effort to select the best and most valuable land available. This would have tended to bias upward the magnitude of potential revenue lost by the grantor, the federal government. Most of the granted land may have been worth considerably more than $1.25 per acre.[5] It seems clear that an orderly, efficient auction system would have generated revenue in the billions of dollars, even at 1870s price levels (see Dennen 1977; Anderson 1985). Given that total government outlays in 1875 were only $274 million, land sales revenue from a policy of maximized revenues could have represented a substantial proportion of annual total receipts.

## Alternatives to Disposal via Land Grants

If the goal of federal land policy was to rid the government of lands held in common, privatization of the public domain was a remarkable

[5]The temporal dimension has largely been neglected in the public domain literature, but obviously the capital value and, hence, the market value will be in part a function of time. Generally the capital value of the public domain—the bulk of which was located in the West—could be expected to increase in real terms as development moved westward. In practice, speculators acquired large areas of land far in advance of development, selling off tracts as development approached and the land's capital value rose. A revenue-maximizing federal land policy would have followed the same pattern, opening up only land on the margin of present development for sale and holding land farther West until development expanded. In this manner, the federal government would have reaped the speculative gains that accrued instead to private speculators.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

success during the 19th century. Between 1796 and 1923, about 860 million acres of federally held land was denationalized, with well over 90 percent of it being transferred to the private sector. As a result, the land area of the public domain in 1912 was only 0.15 percent larger than it had been in 1802, even though the territory of the United States more than tripled over the same period. After 1862, two things had occurred: first, a massive increase in acreage transferred from the public domain to the private sector; second, a gradual decline in the average annual acreage sold and a concomitant decline in annual proceeds. Sales of the public domain went from being a major proportion of total federal revenues before 1840 to the rather trivial 0.33 percent in 1899.

This transition from revenue creation via sales to nonprice (political) "rationing" of the public domain has often been noted, but it has rarely been acknowledged as being paradoxical. It is impossible to know how much revenue the federal government lost because of the land grant programs; much of the land granted was not even surveyed at the time, let alone assessed in any systematic manner.[6] While some portions of the public domain grants might not have been salable at a price high enough to produce revenue net of administrative and surveying costs, only a small percentage of total land grants was "junk land." Overall, hundreds of millions of acres of potentially marketable land were rationed by legislative allocation. In instituting the large-scale system of land grants after 1862 (and to a limited extent before that date), politicians and bureaucrats in effect abandoned hundreds of millions of dollars in revenue.

The bulk of the land grants were simple political transfers. Railroad companies, state governments, and other interest groups lobbied Congress for specific land grants, and Congress mandated particular grants to them. The Homestead Act was the first "uncontrollable" transfer program; like Social Security or AFDC, the number of recipients was not exogenous (that is, determined directly by Congress) but endogenous to the program itself. The major difference in the case of the homestead program was that transfers were in-kind rather than cash.

Some have argued that disposal of the bulk of the public domain by means of land grants was the only feasible method available to Congress due to the high transactions costs of the auction system. Supposedly, preventing squatting and other infringements of prop-

[6]This is especially true in the case of the state and railroad grants, where recipients selected land they were later granted, often after extensive surveys (see Gates 1968, pp. 321–23; Johnson 1976, pp. 68–69).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

erty rights was not economically feasible, land fraud was potentially endemic and difficult to stop, and the means to establish the full value of the land or even to conduct an effective inventory of resources were sorely lacking. A "resort to the ideal auction markets of classical economics" was not "feasible in the actual historical context" (Lebergott 1984, p. 273).

This argument is dubious. During the late 18th and early 19th centuries, the federal government showed that it was capable of efficiently and ruthlessly evicting squatters. Evicting squatters on Indian treaty lands was a major responsibility of U.S. Army units stationed in the West until late in the century.[7] Moreover, it would have been generally unnecessary to have rapidly evicted each squatter following the initial trespass—which obviously would have been expensive—for two simple reasons. First, squatters were not appropriating nonrenewable resources (for example, mineral deposits) but were only farming; trespassers could remain on a tract until it had actually been purchased. Hence, the government would not have to prevent all trespass on the public domain but only to protect nonrenewable resources from theft. Second, laws against trespass probably would have functioned as a deterrent, even if they were imperfectly enforced. (We are ignoring here political constraints on effectively preventing trespass on the public domain because such considerations are beside the point. We are only concerned here with *economic* feasibility.)

The claim that land fraud was an insurmountable problem is dubious for similar reasons. The fraud that has consumed land bureaucrats and public domain historians predominantly involved the evasion of regulatory requirements imposed with respect to land grants; the paradigm of this kind of fraud was the filing of dishonest homesteading claims by the employees of land speculators (Gates 1968, pp. 477–81). Fraud appears to have been only a minor problem with the auction system per se.[8]

Finally, the claim that Congress could not establish the true value of the public domain and that this was a necessary precondition for revenue-maximizing public sales is both contrary to the facts and in violation of elementary economic reasoning. The market value of a given tract could be established by a fair auction; the price paid by the highest bidder would tend to equal the discounted present value

[7]See Lebergott (1984, p. 82). On several occasions the army established a cordon of troops to bar entry by squatters to former Indian land being prepared for settlement (Gates 1968, p. 220).

[8]Far more significant was the overt manipulation of auctions by "claims associations" of squatters who employed violence and threats to keep prices down or prevent auction sales altogether (see Bogue 1963).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of the specific acreage involved. Private land developers would have researched potential purchases and would have had a fairly accurate idea of the market value of any land they bid for. This same reasoning indicates why a thorough governmental survey of the public domain was not a necessary condition for an efficient system of disposal by auction. Private entrepreneurs would have tended to invest time and money to determine tract characteristics prior to the auction in order to estimate probable market value (Gates 1963, pp. 352–53).

Assuming that the auction was a perfectly competitive market, the government would receive the full market value of the auctioned tract even if the seller had no prior relevant information to determine land value. Collusive and other forms of opportunistic behavior would have required careful monitoring, but a thorough survey of the public domain would have been unnecessary. In fact, the land survey system was almost entirely privatized, and the land office contracted with competing private firms for the survey work until very late in the 19th century (Gates 1968, pp. 420–21).

Before the Civil War, proceeds from land auctions—despite the gross, politically mandated leakages built into the system—constituted a major source of federal revenue. And as historian Paul Gates (1968, p. 435) has insisted, for the remainder of the century net revenues from land sales remained fairly constant in real terms, despite the availability of "free" land. But proceeds steadily declined as a percentage of total revenues, and acreage sold was greatly exceeded by acreage simply granted.[9]

## Explaining Land Policy Changes

Explanations for the shift from disposing of the public lands to maximize revenue to granting land outright have been ad hoc at best. Most public domain historians seem convinced that the land grant policy was a boon to the economic development of the frontier and had desirable distributional effects (that is, a form of highly equalitarian resource redistribution); they claim that the system was put

[9]A common assertion was that much western land became valuable only after rail lines were built, permitting development; hence, the federal government was giving up little by granting large tracts to railroad companies (see Shannon 1963, p. 159; Gates 1968, pp. 367–68). The federal government might have leased the land to the railroads through a system of long-term contracts that tied the rental rate on the land to railroad freight income. Such a system would have taken into account the relationship between the value of the land and rail construction, while allowing the Treasury to appropriate some significant portion of the net rent flowing to the railroad companies. Such a simple scheme would have solved the problems supposed by requiring large land grants but would also have generated substantial revenues.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

into effect in large part as a result of an enlightened concern by Congress to serve the public interest by promoting development and democracy (Gates 1984, p. 42). Others have taken a more cynical view, arguing that land policy was formulated by a venal Congress trading the public domain for votes, or that congressmen accepted outright bribes from pressure groups, most notably the railroads and regional interests that expected to benefit from rapid westward expansion (see LeDuc 1966, p. 306). These critics portray the Homestead Act as the outcome of a battle in Congress between the West, which favored the land, and the East, which favored revenue maximization (see Kroos 1966, pp. 77–78). But why would Congress, whether motivated by corruption or public spiritedness, prefer to give away land of unknown or poorly estimated value instead of promoting an efficiently organized system of auction sales?[10]

Western politicians lobbied hard for land grants to western states and surely supported measures like the Homestead Act, which would allow them to expand their tax base, gain popular support among the working classes by advocating "free" land, and promote benefits for special interest groups. These are all elements in an explanation of the *demand* function for the transition from privatization via sales to transfers via grants of land. It is our purpose to suggest a *supply* function for federal land policy that is consistent with the modern economic theory of government and that is at the same time compatible with the main thrust of the argument criticized above.

Our explanation consists of two major elements: first, the size and scope of federal government transfer mechanisms available to Congress for distributing or redistributing revenues to beneficiary groups; second, the role of land-sale proceeds in the context of the federal revenue system.

Consider the modern organization of the federal government with agencies such as HHS, HUD, and SSA, which distribute cash transfer payments to individuals in a bureaucratically efficient manner, or agencies such as the Department of Agriculture, which administer in-kind transfer programs. These organizations and their hundreds of thousands of career employees represent a capital infrastructure that minimizes the transactions costs of transfers to politically favored groups or individuals and prevents payments from being dissipated by flowing to the wrong recipients. These large bureaucracies direct the flow of transfer payments in a politically efficient manner. Viewed as simple vote-maximizers, it is in the interest of members of Con-

[10]For differing views of the question, see Fogel and Rutner (1972) and Dennen (1977).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

gress to ensure that transfers to favored groups are efficiently directed to the intended beneficiaries.

If this transfer capital infrastructure were eliminated, holding all other variables constant, congressional spending would not be directly constrained. Congress could easily throw money out of helicopters or from the roofs of tall buildings, but it would no longer have an efficient mechanism for ensuring that spending flowed into politically determined channels. Despite widely publicized fraud, waste, and abuse, federal agencies *do* effectively channel a high percentage of program spending to politically determined beneficiary groups.[11]

This capital infrastructure has been built up over many years; about 63 percent of federal domestic outlays in 1984 were spent by agencies that did not exist in 1935.[12] The infrastructure represents a large fixed investment that presumably could not be replaced without a long lag if it were eliminated overnight. In any event, the level of income transfer infrastructure during the 19th century was practically zero. Congress had no mechanism to efficiently transfer spending to specific interest groups. Consequently, interest groups did not lobby for cash transfers, but lobbied instead for protection rents (that is, as derived from tariffs) and other substitutes for cash transfers that did not require a large bureaucratic apparatus to administer. Railroad corporations lobbied for choice land grant properties for the express purpose of generating revenues through sale for themselves; they did not lobby for cash subsidies.[13]

It was easy for Congress to monitor the transfer of land in order to ensure that the intended recipient actually took possession. Similarly, while a cash subsidy program to individuals who agreed to settle in the West would have required a large bureaucracy to efficiently administer it, a homestead program could accomplish the same political end at lower administrative cost. At the same time, transfers of land may have been more politically efficient than transfers of cash. At best, the cash value of land resources was likely to be poorly known to potential opponents; hence, the cost of the transfer

[11]Interest groups allocate resources to lobby in Congress for income transfers knowing that they can fully capture the returns on their investment. This is only true given the large transfer infrastructure of the federal government, which in effect protects the property rights in transfers of interest group members.

[12]Calculated from *Historical Tables, Fiscal Year 1986*, Table 5-3.

[13]During the late 1850s, "lobbyists in Washington were running wild in their efforts to secure railroad grants" (Gates 1968, p. 20). The railroads proceeded to set up shop as large-scale realtors, complete with elaborate advertising campaigns. Gates (1934, p. 187) reproduces an 1860 newspaper ad run by the Illinois Central Railroad emphasizing the supposed bargains to be had on 1.2 million acres to be sold for $6.00 to $25.00 per acre. On the promotion of westward expansion by railroads, see Hedges (1928).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

was hard to estimate and could be convincingly alleged to be extremely low. It would have been more difficult to mobilize opposition to such transfers if it was hard to estimate the real cost yet easy for politicians to claim that the cost was low. In fact, throughout the 19th century criticism of "excessive" governmental spending was intense and persistent.[14]

Thus, due to information difficulties associated with the capital valuation of land and intense, widespread opposition to increases in governmental spending, if the public domain had been sold via auction, for political reasons much of the resulting revenue would not have been transferable in the form of cash subsidies. In political terms, land grants may have been the least costly transfer mechanism available. Hence, because of the absence of a bureaucracy capable of administering large cash transfer programs and the likely cost differential between cash and non-cash transfer programs, land grants may have allowed Congress to transfer more real income than under a system of cash transfers. Any increase in revenues from land sales would have tended to function as a substitute for revenue from other sources, causing a corresponding reduction in that revenue.

Recognizing the land grant system as a non-cash federal spending program that transfers direct resources should cause scholars to reassess a peculiar phenomenon of the late 19th century: the "problem" of the federal surplus. The controversy over the federal budget centered on the large surpluses that were run every year from 1866 to 1893; during the 1880s, surpluses amounted to almost 40 percent of federal expenditures (James 1984, p. 193). This behavior by federal politicians is not only grossly inconsistent with modern political proclivities for large deficits, but it casts doubt on modern economic models that postulate a systematic preference for deficits by relatively shortsighted political decision makers (Buchanan and Lee 1982).

The largest single federal spending program of the 19th century, however, was *off budget*. By conservative estimate, Congress freely dispensed nearly 252.6 million acres of the public domain between 1866 and 1893.[15] Over the same period, the total federal surpluses

[14]For public attitudes on the growth of governmental spending before and after the Civil War, see Kimmel (1958, pp. 55, 139–40). Most 19th century Presidents are on record as having expressed concern that governmental spending was growing excessively (pp. 7–55, 61–81). At the same time, the growth rate of governmental spending in real terms was very low; from 1800 to 1860 and again after the Civil War to 1890, real governmental spending grew at an average annual rate of less than 1 percent. See *Historical Statistics of the United States: Part 1*, p. 1049; and *Part 2*, p. 1104.

[15]Totals have been calculated from Hibbard (1924).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

equaled almost $2 billion. If the acreage granted had been sold for an average price of $7.53 per acre, the total proceeds would have equaled the total surpluses.[16] We know that former public domain land in Illinois sold for an average of $9.42 in 1870 and that similar land in Nebraska sold for $11.32 in 1880 (Bogue and Bogue 1963, pp. 378, 384); and millions of acres of railroad grant land were sold by railroad companies for between $4.76 and $11.70 per acre. It is plausible to assert, then, that the federal politicians and bureaucrats were granting land that had an average per acre capital value of at least $7.53. Had the true cost of this off-budget spending program been made a budget item, the large "surpluses" during the 19th century might have been substantially reduced or even converted into significant deficits.

The second major supply-side component focuses on the role of land sale proceeds in the context of the overall federal revenue system, specifically, the interaction between revenues from public land sales and tariffs. Revenues from tariffs constituted the major portion of federal receipts throughout the 19th century. Before the Civil War, proceeds from land sales and tariff revenues were the two major components in federal receipts. The proceeds from these different sources were highly substitutable; one dollar of revenue from land sales could replace one dollar from a tariff and vice versa. There is strong evidence to suggest that this substitutability may have been a significant factor in the demise of the system of revenue-maximizing land sales.

Tariffs represented an important source of protection rents to various domestic producer groups as well as a revenue source for the federal government. Congress enacted and maintained the tariff system because it represented an effective means for achieving both of these objectives. From early in the 19th century, domestic producer pressure-groups and members of Congress promoting their own interests combined arguments for protection with arguments that tariff duties were the only feasible method for raising federal revenue (Pincus 1977, p. 72). In fact, the federal government had limited options in terms of raising revenue; the income tax was unconstitutional and excise taxes were at best marginal revenue generators. Under the circumstances, tariffs were probably the most revenue-productive kind of tax available to the national government. Given this, industry groups competed in Congress to affect the composition of tariffs and their rate schedules (Pincus 1977, 47–66).

[16]Calculated from data in *Historical Statistics of the United States, Part* 2 (1975, p. 1104).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Early in the century, however, the federal government had established another highly productive, though non-tax, revenue-generating device—land sales. Even sales at the minimum price of $1.25 per acre produced comparatively large net revenues. According to a Report of the Commissioner of the General Land Office dated November 30, 1853, the estimated average cost per acre of public lands at time of sale was only 22 cents per acre. By 1836, land sales equaled almost half of total federal revenues.

Because the growth in land sales proceeds was viewed as potentially jeopardizing the tariff levels, the situation concerned supporters of high protective tariffs. Explicit recognition of this threat to continued high tariff levels appears to have first surfaced in the early 1830s in congressional debate over the bill that eventually passed as the Compromise Tariff of 1833 (Sanborn 1900, p. 25). This concern had not emerged earlier because for the first two decades of the 19th century the public debt had remained at high levels and the revenue from land sales was earmarked for debt repayment.[17] The debt was steadily reduced, however, until by the mid 1830s it had been eliminated.

During the late 1820s and early 1830s, members of Congress representing the Northeast, an area of concentrated manufacturing interests favoring high protective tariffs, recognized that the "connection between the public land and tariff was very close . . . [and that] they would soon be confronted with the problem of a surplus" (Stephenson 1967, p. 27). As public debt retirement neared, public land revenues became available to potentially replace tariff revenues. After 1836, pledges made in earlier acts to dedicate land proceeds to repayment of the public debt were effectively abandoned, and such revenues became more substitutable for tariff revenues.

In 1845, Representative Thomasson of Kentucky introduced a bill making grants of 40-acre plots available to settlers and "very frankly stated that one of his chief objects was to remove the public-land fund from the national treasury, as he did not wish a revenue from the lands sufficient to give an excuse for breaking down the protective system" (Sanborn 1900, p. 27). The tariff question again appeared in debate over homestead grants during the early 1850s. In 1850 and 1852, critics charged supporters with favoring a homestead act as a device for maintaining high protective tariffs, apparently an accurate

[17]James (1984, p. 210) argues that the contribution of debt retirement to national income between 1866 and 1890 exceeded that due to the expansion of rail transport. However, most of the surplus revenue used to retire the debt came from high protective tariffs that produced large dead-weight costs, so on balance it is unclear how much of a boon debt retirement was to economic growth.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

description of the motives of many proponents (Sanborn 1900, p. 29). In 1853, Commissioner of the General Land Office John Wilson maintained that income derived from the sale of the public lands had permitted lower tariff rates; he estimated the amount of replaced tariff revenue to that date as being nearly $113.7 million (Gates 1968, p. 22).

Practically speaking, however, there was no simple bifurcation in Congress between supporters of high protective tariffs and supporters of revenue-maximizing sales of the public domain. The principal congressional battles over the public domain from the 1830s to 1862 were between proponents of different competing plans for replacement of the system of land sales as a federal revenue source. Reduced to its simplest elements, the conflict was over plans for ceding the public domain to states within whose boundaries particular lands were found and plans to maintain the system of sales but to distribute the proceeds to all states according to population. The western states preferred the first plan while the eastern states preferred the second.

By the 1840s, the movement for unlimited free land, which had begun early in the century, had become a significant political factor. In 1848, various bills were introduced in Congress that would have made grants to settlers, but they failed to receive much consideration (Sanborn 1900, p. 28). Each year from 1851 to 1854, the House passed a homestead bill and the Senate defeated it (Sanborn 1900, p. 31). Both houses of Congress finally passed a homestead act in 1860. President James Buchanan vetoed it, arguing that the act was an unconstitutional extension of federal powers and that it was unjust to western land holders because the value of their property could be expected to decline (Sanborn 1900, p. 35). In the 1860 election, the Republicans took up the homestead act as a major campaign issue. It finally became law in May 1862 (Gates 1968, pp. 393–95).

Coinciding with the rapid increase in land grants to homesteaders, railroads, and the states after 1862, the federal revenue derived from land sales fell rapidly as a proportion of total receipts. Further, the general decline in tariff rates that had occurred until the Civil War was reversed, and tariff rates began to rise rapidly. Import duty rates, which had reached their lowest level in the century in 1857, increased sharply during the Civil War and remained high for the remainder of the century (Baack and Ray 1983, p. 73). Tariffs continued to be the single most important source of federal revenue after the war ended.

## Conclusion

Federal government land policy began to privatize the public domain during the early 1800s by selling land at auction. Despite inefficien-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

cies, these land sales produced a large proportion of total federal revenues before 1860. Although sales of land continued into the 20th century, after 1862 most of the public domain that was transferred to the private sector was in the form of land grants. We have explained this transition in a manner consistent with the literature on public sector behavior, in which government is modeled as a revenue-maximizing organization that produces wealth transfers to competing pressure groups.

The hypothesis we advance is that the land grant system was a governmental *spending* program. Institutional constraints prevented a large cash subsidy transfer system from emerging; a land transfer system was feasible, however, and was the "second best" solution adopted. To maximize the quantity of resources available for transfer, the government would have had to sell the land in a revenue-maximizing manner and then redistribute those revenues as it became institutionally feasible. Ordinary, real-world, shortsighted congressional decision makers, however, had no choice but to "use it or lose it," as revenues derived from excessive sales were unavailable for cash subsidy transfers in the near term. Consequently, land grants were the major transfer spending program during 19th century.

A further implication of the high marginal costs public sector actors faced in their attempts to use public domain proceeds to increase cash transfers to interest groups was that these proceeds tended to "crowd out" revenues from tariffs. Given the shortsighted perspective of governmental decision makers, reduced tariff revenues harmed an identifiable constituency while increased land sale revenues could not be captured for the benefit of an alternative special interest. In fact, land sales were opposed by the same organized interest groups that were the beneficiaries of protection rents from import duties.

Even though the disposal of the public domain evolved into a massive political transfer program dominated by rent seeking, the outcome may have contributed to economic efficiency. The land grant system did transfer huge acreages from bureaucratic to private ownership, encouraging private individuals to promote "the 'release of energy' by developing and exploiting the vast natural resources of the western United States" (Anderson and Hill 1980, p. 55). We have not attempted to argue that an efficiently organized revenue-maximizing government land sales system would have necessarily directly increased the efficiency of resource allocation, but the probable indirect efficiency gains that might have resulted from an efficient sales system should not be neglected. Assuming that the counterfactual case in which expanded land sale proceeds permitted large reductions in tariff revenues and tariff rates, the deadweight loss due to



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

high protective tariffs might have been significantly reduced and society would have received a net gain in welfare. If we assume that privatization per se enhanced efficiency, then the institutional structure of the privatizing process had important consequences.

Finally, why did the public domain become "locked up" after the late 19th century? Beginning in the 1880s, millions of acres were removed from "private entry" for national parks, wilderness preserves, national resource areas, and so forth. By the mid 1920s, privatization of the public domain had slowed to a trickle. In 1980, the public domain still consisted or over 648 million acres of land. Conventional wisdom holds that federal policy makers became increasingly enlightened with respect to the need to conserve the national heritage. An alternative explanation, consistent with the modern interest group theory of governmental behavior, might focus on two factors: first, the rapid expansion of the large public domain bureaucracies, which constituted a significant pressure group opposed to privatization as it tended to reduce their bureaucratic domain and authority; second, the increased political pressure by various western interests to restrict entry by potential competitors who might reduce the capital value of their property. This shift away from privatization of the public domain presents a potentially fruitful opportunity for future research.

## References

Anderson, Terry L. "TANSTAFL: Rethinking America's Early Land Policy." Political Economy Research Center Working Paper, Bozeman, Montana, 1985.

Anderson, Terry L., and Hill, Peter J. *The Birth of a Transfer Society*. Stanford, Calif.: Hoover Institution Press, 1980.

Baack, Bennett D., and Ray, Edward J. "The Political Economy of Tariff Policy: A Case Study of the United States. *Explorations in Economic History* 20 (1983): 73–93.

Bogue, Allan G. "The Iowa Claims Clubs: Symbol and Substance." In *The Public Lands*, pp. 47–70. Edited by Vernon Cartenson. Madison: University of Wisconsin Press, 1963.

Bogue, Allan G., and Bogue, Margaret Beattie. "Profits and the Frontier Land Speculator." In *The Public Lands*, pp. 369–94. Edited by Vernon Cartenson. Madison: University of Wisconsin Press, 1963.

Buchanan, James M., and Lee, Dwight R. "Politics, Time, and the Laffer Curve." *Journal of Political Economy* 90 (August 1982): 816–19.

Danhof, Clarence H. "Farm-Making Costs and the 'Safety Value': 1850–1860." In *The Public Lands*, pp. 253–96. Edited by Vernon Cartenson. Madison: University of Wisconsin Press, 1963.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Dennen, R. Taylor. "Some Efficiency Effects of Nineteenth Century Federal Land Policy: A Dynamic Analysis." *Agricultural History* 37 (December 1977): 718–36.

Donaldson, Thomas C. *The Public Domain: History with Statistics.* Washington, D.C.: Government Printing Office, 1884.

Fogel, Robert W., and Rutner, Jack L. "The Efficiency Effects of Federal Land Policy, 1850–1900: A Report of Some Provisional Findings." In *The Dimensions of Quantitative Research in History,* pp. 391–93. Edited by William O. Aydelotte, Allan G. Bogue, and Robert W. Fogel. Princeton, N.J.: Princeton University Press, 1972.

Gates, Paul W. *The Illinois Central Railroad and Its Colonization Work.* Cambridge, Mass.: Harvard University Press, 1934.

Gates, Paul W. "The Role of the Land Speculator in Western Development." In *The Public Lands,* pp. 349–68. Edited by Vernon Cartenson. Madison: University of Wisconsin Press, 1963.

Gates, Paul W. *History of Public Land Law Development.* Washington, D.C.: Public Land Law Review Commission, 1968.

Gates, Paul W. "The Federal Lands—Why We Retained Them." In *Rethinking the Federal Lands,* pp. 35–60. Edited by Sterling Brubaker. Washington, D.C.: Resources for the Future, 1984.

Hedges, James B. "Promotion of Immigration to the Pacific Northwest by the Railroads." *Mississippi Valley Historical Review* 15 (September 1928): 183–203.

Hibbard, Benjamin H. *A History of the Public Land Policies.* New York: Macmillan, 1924.

James, John A. "Public Debt Management Policy and Nineteenth-Century American Economic Growth." *Explorations in Economic History* 21 (1984): 192–217.

Johnson, Hildegard B. *Order Upon the Land.* New York: Oxford University Press, 1976.

Kimmel, Lewis H. *Federal Budget and Fiscal Policy, 1789–1958.* Washington, D.C.: The Brookings Institution, 1958.

Kroos, Herman E. *American Economic Development.* Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Lebergott, Stanley. *The Americans: An Economic Record.* New York: W. W. Norton, 1984.

LeDuc, Thomas. "History and Appraisal of U.S. Land Policy to 1862." In *Views of American Economic Growth: The Agricultural Era,* pp. 299–313. Edited by Thomas C. Cochran and Thomas B. Brewer. New York: McGraw-Hill, 1966.

McCormick, Robert, and Tollison, Robert D. *Politicians, Legislation and the Economy.* Boston: Martinus Nijhoff, 1981.

Peltzman, Sam. "Toward a More General Theory of Regulation." *Journal of Law and Economics* 19 (August 1976): 211–40.

Pincus, Jonathan J. *Pressure Groups and Politics in Antebellum Tariffs.* New York: Columbia University Press, 1977.

Rohrbough, Malcolm. *The Land Office Business: The Settlement and Administration of American Public Lands, 1789–1837.* New York: Oxford University Press, 1968.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Sanborn, J. B., "Some Political Aspects of Homestead Legislation." *American Historical Review* 6 (October 1900): 19–37.

Shannon, Fred A. "Comment on 'The Railroad Land Grant Legend." In *The Public Lands*, pp. 157–58. Edited by Vernon Cartenson. Madison: University of Wisconsin Press, 1963.

Stephenson, George M. *The Political History of the Public Lands from 1840 to 1862*. New York: Russell and Russell, 1967.

Stigler, George J. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2 (Spring 1971): 3–21.

Tullock, Gordon. "The Social Cost of Reducing Social Costs." In *Managing the Commons*, pp. 00–00. Edited by Garrett Hardin and John Baden. San Francisco: W. H. Freeman and Company, 1977.

U.S. Bureau of the Census. *Historical Statistics of the United States: Parts 1 and 2*. Washington, D.C.: Government Printing Office, 1975.

U.S. Bureau of Land Management. *Public Land Statistics: 1983*. Washington, D.C.: Government Printing Office, 1983.

U.S. Office of Management and Budget. *Historical Tables: Budget of the United States Government, Fiscal Year 1986*. Washington, D.C.: Government Printing Office, 1985.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Don't Revise the Clayton Act, Scrap It!

*William F. Shughart II*

Early in 1986, word began to leak out that the Reagan administration was developing a proposal to modify the language of the Clayton Act, one of the pillars of U.S. antitrust law. Accounts published in February suggested that the plan has several basic elements (see, for example, Cohodas 1986). One is to change the litmus test for whether or not a merger violates the law from a finding that the effect of the acquisition "may be substantially to lessen competition" to one of a "significant probability" of an anticompetitive effect. Other proposals would require the antitrust agencies to analyze proposed takeovers within the framework of the current Department of Justice (DOJ) merger guidelines, exempt import-injured industries from the antimerger law for a period of five years, place restrictions on the award of treble damages, and relax somewhat the prohibitions against interlocking directorates. Early reports have touted the plan as a major departure from the "big-is-bad" philosophy, which if adopted would bring the antitrust laws up to date, making them more compatible with the reality that in the modern economy U.S. firms must compete in a world market. Assistant Attorney General Ginsburg, for example, has said that "it is not an immodest goal to try to bring the antitrust laws into the 1980s before they're over" (Cohodas 1986, p. 188).

Although the proposal has some merit in the sense that more mergers will presumably go through under the revised language than under the old, these gains are more than offset by the fact that the administration's plan would institutionalize a merger review *process*

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

The author is Associate Professor of Economics and Senior Research Associate at the Center for Study of Public Choice, George Mason University. He was formerly Special Assistant to the Director of the Bureau of Economics at the Federal Trade Commission. The comments of Robert Tollison on an earlier version of this paper are gratefully acknowledged. The usual caveat applies.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

that is based on outmoded economic thinking and, indeed, substantially harms the efficient allocation of productive resources in the economy. To see these perverse effects, let us consider the two main components of antimerger law enforcement: the DOJ merger guidelines and the Hart-Scott-Rodino premerger notification requirements.

## The Reagan Merger Guidelines

The DOJ merger guidelines, first promulgated in 1968 (1 Trade Reg. Rep., Par. 4510, May 1968), establish the analytical model that the government applies in evaluating proposed mergers. The model is unapologetically structuralist, that is, based on the notion that the degree to which competition prevails in a market is directly linked to the number and size distribution of firms in the industry. Thus, whenever a merger is proposed, the first and in many respects the most important exercise is to calculate a numerical index of market concentration. Under the 1968 guidelines, the crucial number was the four-firm concentration ratio, or the share of sales accounted for by the four largest firms in the industry. The revised guidelines issued in 1982 replaced this measure with the Hirschman-Herfindahl Index, which is calculated by summing the squares of the individual market shares of all industry members.[1] In either case, however, mergers proposed in "concentrated" industries—so defined when the numerical concentration measure exceeds a preestablished critical value—are more likely than not to be challenged by the antitrust authorities.

The notion that industry structure represents an important source of market power stands in stark contrast to the conclusions one can draw from the empirical evidence that has been accumulating since the early 1970s. The studies of Brozen (1971), Demsetz (1971), Peltzman (1977), and others suggest that the above-normal profit levels observed in highly concentrated industries are driven by economic efficiency rather than by the exercise of market power.[2] That is, the data imply that high levels of industrial concentration can be better explained by the survival over time of a small number of efficient firms whose profitability is traceable to an ability to economize on the use of productive resources. Profits are high not because prices are high but because costs are low. This of course means that a social policy hostile to high concentration operates to the detriment of consumer welfare.

[1]Subsequent, but unimportant, guideline revisions were made in 1984.
[2]Disbelievers may wish to consult Scherer (1979).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Although the new merger guidelines pay much more attention than did their predecessor to efficiency considerations, technological change, foreign competition, and other factors mitigating the potential anticompetitive effects of merger, they retain the numerical concentration measures that are the meat and potatoes of structural antitrust analysis. More importantly, the structural numerics were not made very much looser. To illustrate the difference between the two sets of merger criteria, consider a merger in a "concentrated" market assuming that the acquiring firm accounts for 15 percent of industry sales. Under the 1968 guidelines, mergers in which the target firm's market share was 1 percent or more would likely be challenged by the antitrust authorities. The threshold market share for an acquired firm in similar circumstances was raised to about 5 percent under the new guidelines. Thus, the screen is now looser, but not by much.

The failure of the Reagan guidelines to relax substantially, if not eliminate, numerical merger criteria is important not because rules for evaluating mergers are by themselves undesirable, but because such criteria preserve the essentials of the structuralist approach to antitrust policy.[3] Given a modicum of creativity on the part of antitrust lawyers and economists, markets can be defined in such a way that any industry becomes "concentrated." For example, Franklin Fisher (1979) reports that in challenging Nestle's acquisition of Stouffer's, the Federal Trade Commission (FTC) defined the relevant market to be that for "high-priced, nonethnic, frozen entrees." Thus, even though more mergers presumably pass muster under the new criteria than the old, the new guidelines do little to change the process of merger analysis. Evidence to this effect is given by the fact that the antitrust bureaucracy took to the new guidelines with remarkable alacrity.

The Reagan administration proposal to codify the merger guidelines in the language of the Clayton Act therefore comes as a major disappointment. If adopted, the plan will at best substitute a "big-is-probably-bad" philosophy for the "big-is-bad" approach that has guided antitrust enforcement for decades. Business as usual will prevail at the antitrust bureaus and, what is worse, future administrations, perhaps bringing with them an inclination toward less government intervention in the affairs of the private economy, will find it more difficult to undercut the enforcement agencies' penchant for discovering monopoly on every street corner.

[3]Another basic issue is whether the guidelines are really guidelines at all. One study has shown that over 20 percent of the mergers challenged under the 1968 guidelines fell *below* the threshold criteria. See Rogowsky (1984).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

What explains the administration's apparent acceptance of the structural approach to antitrust policy, that is, who gains from narrow merger guidelines? The answer seems clear. Bureaucratic incentives run strongly in the direction of producing visible output, and tighter screens give the enforcement agencies more mergers to investigate. The more work there is for government, the more opportunities there are for the attorney staff to build the human capital that is rewarded when they subsequently take jobs in big antitrust law firms, and the larger and more secure are the antitrust bureaus. The private antitrust bar gets to defend more clients, and economists working as consultants on both sides of the issue earn larger incomes. Thus, the private interests of those employed in the antitrust industry overcame economics in preserving the status quo.

An interesting sidelight on the political economy of antitrust enforcement is provided by the FTC's reaction to the DOJ's revised merger guidelines. On the same day that the 1982 guidelines were promulgated, the FTC, for the first time in its history, issued a statement of "principles" concerning horizontal mergers (see Federal Trade Commission 1982). Although the staffs of the two agencies had coordinated their efforts while the documents were in preparation and, indeed, Chairman Miller (1982, p. 2) emphasized the "harmony" between the agencies' enforcement philosophies, the FTC statement differed from the DOJ guidelines in an important way, namely, it contained no numerical merger criteria. Chairman Miller justified this approach on the ground "that it would be unwise for the two agencies to publish fairly precise standards that might differ in even minor respects." Thus, although the FTC (1982, p. 2) argued that the weight of economic evidence justified some "revision of market share benchmarks," it chose not to adopt numerical merger criteria so as to avoid any possible confusion between FTC and DOJ standards.

The consequence of this difference was quite dramatic—the FTC's statement was virtually dead on arrival as a matter of public merger policy. Attorneys for the FTC quickly embraced the DOJ guidelines because in contrast to the approach taken by their own agency, the DOJ standards were quite compatible with traditional merger analysis. Quantitative concentration data are much more understandable than the non-numeric, non-market-share considerations contained in the FTC's statement. Thus, although the revised DOJ guidelines required that users learn the arithmetic of calculating a new index of market concentration, the DOJ approach dominated because it was mechanical. Merger analysis could continue to focus on the issue of market definition and market shares.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

## Premerger Notification

The so-called Hart-Scott-Rodino (H-S-R) premerger notification process, established by the Antitrust Improvements Act of 1976, requires firms above a given size to announce in advance their intention to merge. Designed to give the antitrust enforcement agencies adequate time to assess the competitive effects of proposed takeovers prior to their consummation, the H-S-R rules also penalize firms that have a comparative advantage in discovering undervalued assets in the economy. The penalty is especially heavy when the proposed merger involves two firms in the same industry.

The H-S-R process establishes the following timetable. The acquiring firm first files premerger notification forms simultaneously with the FTC and DOJ that announce to the government its intention to merge and contain preliminary information about the parties involved in the acquisition. The two enforcement agencies then have 10 days to decide which of the bureaus is to have jurisdiction and whether or not to seek additional information from the prospective merger partners. If no such "first request" is forthcoming, the merger may proceed. Assuming that more data are sought, however, a second decision point is scheduled after the firms have complied. This involves whether or not to issue a "second request" for information about the acquisition. Again, the merger can go forward if the agency does not require the parties to produce additional data. But if so, once the second request has been fulfilled the government must act within a specified time limit (10 days for a cash deal and 20 days if the acquisition involves the purchase of stock) if it decides to challenge the takeover. It is worth pointing out that the paperwork burden of second requests can be immense. On occasion, firms have literally been required to produce documents in truckload lots. It is also worth noting that because the time limits in which the government must act to prevent a merger from going forward do not commence until the antitrust agency is satisfied that the parties have complied with information requests, in practice the steps in the review process are measured in months rather than days.

Premerger notification essentially requires firms to announce publicly that they have discovered a profit opportunity in the economy. In cases where the acquisition is postponed while the government seeks additional information from the prospective merger partners, other firms, which had been unaware of the existence of undervalued assets, are given time to step forward with takeover offers of their own. The H-S-R process thus allows these other firms to free ride on the information revealed by the premerger announcement. This tends



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

to reduce the efficiency with which resources are allocated in the economy for two reasons. First, it lowers the value of information about the existence of profit opportunities to firms operating in the same industry.[4] Second, it promotes conglomerate mergers because such takeovers are less likely to be delayed by the H-S-R process. Both of these consequences are costly to the economy because firms are more likely to discover undervalued assets within their own industry and to employ such resources more efficiently when purchased than companies operating in an unrelated line of business.

Those who want tougher antimerger law enforcement often complain that the administration's policy lets through too many mergers that are too big and do not work. Yet, this is precisely the consequence of combining premerger notification with a narrow set of merger guidelines. It is not by accident that in recent years steel companies have acquired oil companies, or that tobacco companies have merged with soft drink manufacturers. Rather, such acquisitions have been forced on the economy because the firm that typically steps forward when a previously announced merger falls through due to antitrust considerations is one that represents no overlapping markets. The critics cannot have it both ways.

The "cola wars" of 1986 illustrate another cost of premerger notification. Shortly after Pepsico made public its intention to acquire the assets of Seven-Up from Philip Morris, Coca-Cola filed premerger documents announcing an agreement to purchase Dr. Pepper. Subsequent newspaper reports suggested—and a federal district court judge held in granting a preliminary injunction against Coca-Cola's planned acquisition—that the merger announcement had been a sham designed to plant in the minds of the FTC the idea that if it did not challenge the Pepsi–Seven-Up deal, it would be faced with a merger wave in the soft drink industry. Regardless of the merits of either acquisition, it is clear that the disclosure requirements and delay involved provide the opportunity for firms to use the H-S-R process as a competitive weapon.

## Conclusion

Under the guise of helping U.S. firms compete in the world market, the Reagan administration proposal to revise the Clayton Act will have the unfortunate consequence of enshrining a merger review process that is out of step with current economic research and, in

[4]In this regard, premerger notification is likely to generate adverse capital market effects similar to those associated with the disclosure requirements imposed on cash takeover bids by the Williams Act. See Jarrell and Bradley (1980).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

fact, raises substantial barriers to the efficient allocation of productive resources in the economy. Given the conservative-libertarian orientation of the President's appointees to the antitrust agencies, the plan is nothing short of dismaying.

What is the alternative? A modest proposal with a great deal of merit would allow the market rather than the government to evaluate the pro- and anticompetitive effects of merger. That is, consumers would be much better off with a loose merger screen that permitted all but the most overwhelmingly large takeovers to go forward without government interference. The antitrust authorities could then observe the actual as opposed to the predicted effects, and take action to dissolve the merger when appropriate. More efficiency-enhancing mergers would take place in the economy and antitrust law could be invoked ex post with hard evidence in hand rather than ex ante on the basis of "significant probabilities" and the complaints of competitors. If our only choice is between the administration's proposal and doing nothing, however, it would be far better to leave the Clayton Act alone. By not giving the merger guidelines the force of law, the latter alternative is at least less likely to prevent future administrations from taking bolder steps to truly modernize public policy toward merger.

## References

Brozen, Yale. "Bain's Concentration and Rates of Return Revisited." *Journal of Law and Economics* 14 (October 1971): 351–96.

Cohodas, Nadine. "Reagan Seeks Relaxation of Antitrust Laws." *Congressional Quarterly Weekly Report* 44 (1 February 1986): 187–92.

Demsetz, Harold. "Two Systems of Belief about Monopoly." In *Industrial Concentration: The New Learning*, pp. 164–84. Edited by H. J. Goldschmid, H. M. Mann, and J. F. Weston. Boston: Little Brown, 1971.

Federal Trade Commission (FTC). *Statement of Federal Trade Commission concerning Horizontal Mergers*. Washington, D.C., 14 July 1982.

Fisher, Franklin. "Diagnosing Monopoly." *Quarterly Review of Economics and Business* 19 (Summer 1979): 7–33.

Jarrell, Gregg, and Bradley, Michael. "The Economic Effects of Federal and State Regulations of Cash Tender Offers." *Journal of Law and Economics* 23 (April 1980): 371–407.

Miller, James C. III. *Prepared Statement of James C. Miller III, Chairman, Federal Trade Commission, on Merger Enforcement Policy, Submitted to the Committee on the Judiciary, United States Senate*, 21 July 1982.

Peltzman, Sam. "The Gains and Losses from Industrial Concentration." *Journal of Law and Economics* 20 (October 1977): 229–63.

Rogowsky, Robert A. "The Justice Department's Merger Guidelines: A Study in the Application of the Rule." In *Research in Law and Economics*, vol.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

6, pp. 135–66. Edited by Richard O. Zerbe, Jr. Greenwich, Conn.: JAI Press, 1984.
Scherer, F. M. "The Causes and Consequences of Rising Industrial Concentration." *Journal of Law and Economics* 22 (April 1979): 191–208.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Labor Law and Labor-Management Cooperation: Two Incompatible Views

*Charles W. Baird*

## Introduction

There are few informed observers of the labor-management scene who do not agree that U.S. labor relations and its institutions need fundamental reform. In two earlier papers (Baird 1985b, 1987), I considered some of the ways that current American labor law, especially the National Labor Relations Act (NLRA), adversely affects the outcomes of the entrepreneurial discovery process in labor-management relations, and thus impairs the ability of American unionized and nonunionized firms to compete effectively in international markets. In June 1986, Stephen Schlossberg, Deputy Under Secretary in the U.S. Department of Labor (DOL), released a report he coauthored with Steven Fetter, also of DOL, concerning the potential conflict between existing labor law and the increasing attempts of management and workers to cooperate in the discovery and development of ways to make American industries more competitive. Schlossberg and Fetter (hereinafter, SF) agree with me that existing labor law, especially the NLRA, can impede the discovery and development of effective forms of worker-management cooperation. However, their diagnosis and suggested remedies are diametrically opposed to my own.

Although SF recognize that existing labor law and practices can impede the discovery and implementation of such innovations as worker participation in management and quality of work-life programs, they insist that those impediments be removed in ways that will maintain and strengthen existing Wagner Act institutions of

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at California State University–Hayward. He thanks his colleague, Nan Maxwell, and Francis A. O'Connell and Howard Dickman for helpful comments on an earlier draft. The usual disclaimers apply.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

compulsory, monopoly bargaining. They invite labor-management practitioners and academicians to identify, develop, and promote changes in labor law that would make it easier for its compulsory aspects—mandatory collective bargaining between exclusive bargaining agents and employers, union security arrangements, job property rights for strikers, and the proscription of company unions—to coexist with newly emerging experiments in cooperative labor-management relations. They even suggest that labor law ought to be changed to make it easier for unions to gain monopoly representation rights over workers in the first place.

In other words, SF see newly emerging forms of worker-management cooperation as a threat to monopoly compulsory bargaining and want to develop changes in the law that will protect and extend monopoly compulsory bargaining. To SF, cooperation is acceptable only if monopoly compulsory bargaining is not threatened. In their view, the end to be achieved is the protection and extension of monopoly compulsory bargaining. To them, existing labor law must be changed to promote the kinds of cooperation that are consistent with that end.

On the other hand, I propose that the entire framework of compulsory unionism be scrapped in favor of a system of voluntary unionism within which there can be no conflict between labor law and labor-management cooperation. I see monopoly compulsory bargaining as a threat to the discovery and development of superior modes of labor-management relations, and I advocate repeal of existing labor laws to eliminate that threat. In my view, the end to be achieved is the discovery and implementation of superior modes of labor-management relations within which cooperation can flourish. To me, existing labor institutions are a threat to that end and must be removed. In the remainder of the paper, I will first outline my arguments and then explain and critique those of SF.

## Government Regulation and Entrepreneurial Discovery

My arguments in favor of deregulation of the labor market are based on the Mises-Hayek-Kirzner analysis of the competitive entrepreneurial market process.[1] The key insight of that analysis is that every situation of discoordination, economic error, or institutional failure contains profit opportunities for those entrepreneurs alert enough to

[1]The best exposition of this analytical framework is Kirzner (1973, 1985). See also Hayek ([1968] 1978).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

notice them and act on them. When entrepreneurs act on such opportunities the unintended consequence of their actions is that the discoordination, error, or failure is ameliorated.

Standard neoclassical criticism of government regulation focuses on two sources of government failure: the knowledge problem and the political problem. Inasmuch as the knowledge necessary for regulators to improve upon market outcomes when there has been some sort of "market failure" exists nowhere in its entirety, but is, rather, divided up in widely dispersed bits, regulators must make their decisions, which are to apply to the whole economy, in ignorance of most of the necessary knowledge.[2] Thus their decisions are likely to make things worse rather than better. Furthermore, regulatory decisions are political decisions. They are inevitably made in response to political pressures, not a dispassionate consideration of costs and benefits. Only some costs and benefits are considered—those that affect interest groups with political clout.

Israel Kirzner (1985, chap. 6) has pointed out a third source of government regulatory failure—the effects of regulation on the outcomes of the entepreneurial discovery process. There are four such effects, which he labels "unsimulated discovery," "undiscovered discovery," "stifled discovery," and "superfluous discovery."

*Unsimulated Discovery*

It is impossible for government functionaries to simulate the entrepreneurial discovery process because (a) the division of knowledge makes it impossible for them to use most of the relevant knowledge and (b) it is difficult for them to capture the profits that result from successful discovery for themselves, so they will not be as alert to such opportunities as they would be if they could.

Even a bureaucrat who is dedicated to the public interest is unlikely to be motivated to discover that which is as yet undreamed of by anyone. The most a dedicated bureaucrat is likely to do is engage in purposive search for that which he knows he does not know. Most bureaucrats spend their time making decisions regarding known alternatives rather than seeking out new opportunities. And bureaucrats are like everyone else—they are dedicated primarily to what is in their own interest. Moreover, in contrast to the entrepreneurial discovery process in the private marketplace, there is no mechanism in government regulation that assures that less competent regulators will be replaced with those who are more competent.

[2]The notion of the division of knowledge, although employed in neoclassical analysis, originated with Hayek ([1945] 1948).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

*Undiscovered Discovery*

It is impossible for government regulators to know where the unregulated market discovery process would, if left alone, lead. Thus, even if some sort of genuine marketplace problem exists, the imposition of regulations based on less knowledge than is discovered and used in the market discovery process may forestall corrective entrepreneurial discoveries that otherwise would have been forthcoming. Regulators should not assume that because the market has not solved a problem now, it is incapable of doing so. The market discovery process is not instantaneous. If regulations are imposed the discoveries that otherwise would have been made may not be made. And genuine solutions, solutions based on more complete knowledge than any regulator or group of regulators could ever have, will never become known and implemented.

*Stifled Discovery*

Government regulation often takes the form of restriction of particular market activities such as market entry by interlopers who may have new ideas about how to do things. A major vehicle for entrepreneurial discovery is the freedom of newcomers to innovate. If gains from innovation by interlopers and those who would try alternative organizational and institutional arrangements are foreclosed by blocked entry (government-granted monopoly) there will be less alertness to such possibilities than there otherwise would be. Discoveries that newcomers might have tried and proved, and others might have adopted, will never be tried, proved, and adopted.

*Superfluous Discovery*

Finally, entrepreneurs try to do the best they can for themselves given the circumstances and constraints they confront. The imposition of government regulations may push entrepreneurial discovery in wholly unanticipated and unwelcome directions. Bureaucrats cannot anticipate the profit opportunities that are opened up by newly imposed regulatory constraints and thus cannot predict where the resulting discovery process will lead.

This is in addition to the neoclassical point that government regulation diverts alertness away from market competition and innovation toward political competition for special favors. General benefits through entrepreneurial discovery in the private market are sacrificed in favor of attempts to discover opportunities for political gains by organized interests who pursue narrow benefits for themselves at the expense of all others. The positive sum game of entrepreneurial



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

discovery in the private market is gradually replaced by the zero sum game of entrepreneurial discovery in the political market.

## Three Labor Laws

The Clayton Act (1914), the Norris-La Guardia Act (1932), and the National Labor Relations Act (1935, amended 1947) have important features that adversely affect the outcomes of entrepreneurial discovery in the labor market. The Clayton Act (CA) was the first law to grant labor unions immunity from prosecution under antitrust law. The Norris-La Guardia Act (NLA) reinforced the union antitrust exemption and outlawed union-free agreements (the so-called "yellow dog" contracts) between employees and employers. The National Labor Relations Act (NLRA), as amended, has four features especially germane to the present inquiry: exclusive representation, union security, the proscription of company unions, and job property rights for strikers.

Under the CA and the NLA, one union can represent all the workers in an entire industry, and unions that represent workers in different industries can affiliate with each other without fear of violating antitrust laws. For example, the United Steel Workers (USW) can represent the workers who are employed by *all* the steel companies, and the United Auto Workers (UAW) can represent *all* the workers who are employed in the automobile industry, and the USW and the UAW can affiliate as members of the AFL-CIO. If all the steel companies, or all the auto companies, got together to collude on price, quality, and marketing strategies in the output markets of those industries, they would run afoul of the Sherman Antitrust Act. But USW and the UAW need fear no legal ramifications if they do the same in the labor markets of those industries.

If interloper labor representation organizations were free to enter the labor representation market, antitrust laws would be superfluous at best and, as in product markets, harmful in fact. But, because of the monopoly-granting features of the NLRA to be discussed below, the labor representation market is not open. Thus the monopoly power of industry-wide unions is not checked by normal market processes. The AFL-CIO, free of fear of antitrust prosecution, prohibits or regulates competition among member unions for representation rights; so competition among AFL-CIO affiliated unions cannot be relied on to keep incumbent exclusive bargaining agents honest. AFL-CIO unions do face organizing competition in some employments from unions outside the federation, such as the Teamsters, but even this competition can be quashed by "gentlemen's agreements"



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

without fear of antitrust prosecution. And existing labor law makes it illegal for independent interloper employee organizations of diverse types, such as employer-sponsored unions, to compete with incumbents.

Section 3 of the NLA provides that the courts may not enforce agreements between employees and employers whereby initial and continued employment is contingent upon an employee's abstension from membership in, or affiliation with, a labor union. Such union-free agreements were frequently requested by *employees* as protection against the threats and arguments of union organizers (Reynolds 1984, pp. 97–100). They were also, of course, frequently used by employers to make it more difficult for unions to organize workers.

Section 9(a) of the NLRA states that a union that is selected by a majority of employees in a bargaining unit of a firm in a union representation election shall be the exclusive bargaining agent for all the employees in that bargaining unit. Employers and dissenting employees cannot prevent such certification elections from taking place if a union gains the signatures of 30 percent or more of the workers asking the union to represent them. The winning union not only gets to represent those employees who freely choose such representation, it also gets to represent those employees who want to be represented by some other union as well as those employees who do not want to be represented by any union. The winning union gets to be a monopolist in the provision of representation services to the employees. Competition from other unions and from nonunion modes of representation, such as employee committees set up to cooperate with management on developing improvements in work rules and productivity, is blocked by force of law.

Section 8(a)3 of the NLRA empowers labor unions that are exclusive bargaining agents to agree with employers that all employees either must join the union shortly after being hired (a union shop) or, in other cases, may refrain from joining the union but still must pay union dues (an agency shop). Not only must employees who do not want to be represented by the union acquiesce to such representation "services," where there is a union shop or an agency shop they must pay for what they do not want or be fired. Not only is the union granted a government protected monopoly, it is empowered to force people to pay for what it monopolistically provides no matter how poorly it performs. These two forms of union security are designed to protect the unions against dissatisfied employees who would otherwise withhold support from unions that perform poorly. One of the most important mechanisms of the competitive market process—



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

escape from poor performance—is blocked.[3] The only escape routes provided by the NLRA are costly and time-consuming decertification and deauthorization campaigns. Individuals cannot escape on their own.

Section 8(a)2 of the NLRA forbids employers to be involved in the "formation or administration" of any "labor organization" or to "contribute financial or other support to it." This feature of the law is addressed to the alleged role of company-sponsored unions in the years prior to the passage of the original Wagner Act. It was widely asserted that such so-called company unions acted merely as fronts for the employer who would use them to exploit employees. According to union folklore, company unions were formed merely to meet the requirements of Section 7(a) of the 1933 National Industrial Recovery Act so that the "real" unions (mainly those affiliated with the American Federation of Labor) could be kept out. The truth of the matter is that most company-sponsored unions were formed in an effort by both employees and employers to discover effective modes of labor management relations. Many of them could reasonably be regarded as early forms of what today are called "quality of work life" arrangements whereby employees participate in the formation of management decisions regarding production arrangements. The company unions stressed the fact that labor and management are complementary members of the production team. The "real" unions—those affiliated with the AFL and, later, the CIO—on the other hand, promoted the view that management and labor are natural enemies. It should come as no surprise that labor-management cooperation is difficult to achieve under the aegis of "real" unions.

The massive recognition strikes of 1933–35 were, by and large, not strikes by a majority of employees of the involved firms. More often than not an outside union got a few employees of a firm to go on strike and then the union would send in "flying squadrons" of nonemployees to act as pickets and to intimidate the majority of the employees, often members of company-sponsored unions or small independent unions not approved by the AFL, who wanted to continue to work (Baird 1984, pp. 42–44). The role of company unions was nicely stated by Judge John P. Nields in his decision in the 1934 *Wierton Steel* case, which involved a fight between a Wierton-sponsored and an AFL union (Baird 1984, p. 43):

[3]For a discussion of the pros and cons of exclusive representation and union security, see Baird (1984, pp. 48–51, 55–59). Of course, under a system of voluntary unionism, with no mandatory bargaining and no legislatively enforced exclusivity, union security arrangements would be consistent with the principles of voluntary exchange contracting.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> It is said that this relation (between management and workers) involves the problem of an economic balance of the power of labor against the power of capital. The theory of a balance of power is based upon the assumption of an inevitable and necessary diversity of interest. This is the traditional old-world theory. It is not the 20th century American theory of that relation as dependent upon mutual interest, understanding, and goodwill. The modern theory is embodied in the Wierton plan of employee organization.

Section 13 of the NLRA affirms the legal "right" of unionized workers to strike. Section 2(3) stipulates that a striking worker maintains his or her status as an employee, and case law, based upon Section 8's proscription of "unfair labor practices," has established that strikers have property rights to their jobs after a strike is settled. Strikers in an unfair labor practice strike must be immediately rehired when the strike is settled. No replacement worker can keep a job if doing so means that a striker would lose a job. In an economic strike strikers have first claim on any jobs vacated by replacement workers who were hired during the strike. But an economic strike can be reclassified into an unfair labor practice strike if an employer is deemed to have failed to bargain "in good faith" over the economic issues. Thus, as a result of the NLRA, a strike is more than just a collective withholding of labor services in face of unacceptable offers of terms of employment. A strike is also a denial of the right of nonstrikers permanently to replace those who strike. And during strikes strikers can, with impunity, through picket lines, attempt to block access to the struck firms of nonstriking workers, suppliers, and customers.

## Labor Law and Entrepreneurial Discovery

It is no secret that many basic industries in the United States such as autos, steel, and textiles are in trouble. They are unable to compete effectively with foreign producers, notably the Japanese, and so they are declining. In response to this decline, both management and unions have appealed to the federal government for protection against foreign competition and for taxpayer subsidies. In addition, unions have appealed for plant closing restrictions. I contend that it is no accident that the industries that are most imperiled by competition from the Japanese are precisely those industries that are most heavily unionized according to the strictures of current labor law.

The CA-NLA exemption of labor unions from antitrust law, in light of the closed market in labor representation services set up and enforced by the NLRA, makes competition between existing labor unions and between existing unions and interlopers virtually impos-



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

sible. If different unions represented the employees of the different firms in the auto industry, for example, competition between them would perhaps have forestalled the establishment of the 83 percent premium of employee compensation in the auto industry over the average for all manufacturing employments in 1985 (Reynolds 1986, p. 112). The development of restrictive work rules that impair entrepreneurial discovery by workers of ways to improve productivity might never have developed. The large compensation premium and the restrictive work rules have not been in the long-run interests of either the employees or the employers. The entrepreneurial discovery process that would otherwise have identified and implemented pay and work-rule arrangements that were mutually beneficial on a long-run basis was stifled, and the industry was the victim of undiscovered discovery.

The NLA's proscription of enforceable union-free agreements between employees and employers cuts off experimentation with guaranteed union-free environments. Experimentation with personnel practices and work rules in nonunion firms is often circumscribed by fears of both employees and employers that union organizers will exploit any perceived problems that may emerge. Under the threat of monopoly unionism it is better to play it safe than to be too innovative.

The four features of the NLRA previously cited—exclusive representation, union security, proscription of company unions, and the creation of job property rights for strikers—have also blocked the discovery of superior forms of labor-management relations that would have permitted American unionized industries better to adapt to changing market conditions. And they have given rise to superfluous discovery that hurts nonunion as well as unionized industries.

Where there is a certified exclusive bargaining agent, there cannot be active competition between two or more labor unions at the same firm at the same time. Neither can there be competition between union and nonunion providers of representation services. Individuals cannot even represent themselves. Protected by its government-granted monopoly, an exclusive bargaining agent has less incentive than it otherwise would have to be alert to possible innovations which would benefit both managers and workers by, for example, lowering costs and thereby allowing the firm to be a better competitor. Indeed, the exclusive bargaining agent has little incentive even to be alert to opportunities to better improve its own services to employees. It does not have to fear that dissatisfied employees will seek representation services elsewhere or decide to represent themselves. The only way individual employees can escape from union abuse is to



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

organize a majority in favor of undertaking a lengthy and costly decertification procedure. If more than one union were permitted to function at the same time, and if individual employees always had the ability immediately to opt out of the services provided by an unsatisfactory union, every union would have a strong incentive to remain alert to opportunities to improve its services to employees.

The attempt to stifle experimentation with alternative forms of labor-management relations under the rubric of exclusive representation is well illustrated by the recently announced agreement between General Motors and the UAW regarding GM's Saturn project to be undertaken in Spring Hill, Tennessee. GM has agreed, in a "prehire agreement," to recognize the UAW as the exclusive bargaining agent for all Saturn employees even though those employes have not yet been identified and hired. The prehire agreement gives hiring preference to UAW members who have been laid off from other GM plants. The Saturn project is an attempt to identify ways for GM to cope more effectively with Japanese auto competition. But even before the project gets underway, GM has forsaken any chance of innovation in labor-management relations by chaining itself to the failed practices of the UAW.

Why would GM do such a thing? I think it is because GM considers itself stuck with the UAW in most of its plants. The monopoly advantages given the UAW by the NLRA have made getting rid of the UAW very difficult. GM figures that if it is stuck with the UAW it is going to make it more difficult for other auto companies, especially Japanese plants set up in the United States, to avoid the UAW. Nissan has set up a plant in Smyrna, Tennessee, 30 miles from the Saturn site. Giving the UAW a toehold in the region makes Nissan more vulnerable. Alfred Warren, GM's vice president for industrial relations, put it this way: "I do not believe the UAW can afford to let Nissan go unorganized. If they continue without a union, other Japanese manufacturers might get the idea that they too can work in the United States without a union."[4]

The principle of union security makes it even easier for exclusive bargaining agents to survive without being alert to opportunities to discover how better to serve the common interests of employees and employers. If employees cannot even withdraw financial support from union officers who perform unsatisfactorily, those officers can tune their alertness to ways to better serve themselves rather than their members. Protected by forced union dues, they are free to undertake what, from the viewpoint of employees and employers,

[4]Quoted by Reynolds (1986, p. 121)



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

are wholly superfluous discovery processes. A symptom of this is the current enthusiasm of labor union leaders for laws enforcing the doctrine called "comparable worth." If such laws were passed, wages would no longer be determined by collective bargaining; they would be determined by political wage boards. Apparently union officers now think they are better at political manipulation than collective bargaining (Baird 1985a).

The proscription of employer involvement in labor unions cuts off an obvious avenue of entrepreneurial discovery in labor relations. Since workers and management are complementary participants in production, it seems likely that if they were free explicitly to cooperate in experimenting with alternatives to the existing structure of unionism, significant innovations would be discovered and implemented. For example, it was a company union, the Leeds and Northrup Cooperative Association, that, in 1922, first introduced and experimented with unemployment insurance. Other company unions—e.g., the Industrial Assembly of Goodyear Tire and Rubber, the Filene Cooperative Association, and the Employees Representation Plan of Colorado Fuel and Iron Co.—were significant contributors to the development of modern personnel programs (Nelson 1982).

The fact that the NLRA prohibits such discovery is well illustrated in *NLRB v. Cabot Carbon Co.* (360 U.S. 203 [1959]). In this case the U.S. Supreme Court held that employer-created employee committees set up to confer with management on a regular basis on matters of mutual concern—such as quality control, production innovation, and complaint settlement—violated Section 8(a)(2) of the NLRA because the committees were, in the language of Section 2(5), "dealing with" employers and thus were "labor organizations." Although three district courts of appeal have subsequently rendered decisions based on a broader interpretation of the definition of a labor organization, the Supreme Court has never modified its ruling in *Cabot Carbon*, and that case is still used by the NLRB and other federal courts to prohibit such forms of labor-management cooperation.

The Japanese model of labor-management relations is often held up as the one we ought to emulate. That model is squarely based on employer participation in labor unions. It is striking to realize that the infamous company unions of the 1930s could have evolved into labor-management institutions similar to those so highly regarded and recommended today. If it had not been for the adoption of the NLRA they might have done just that. Or some altogether different and superior forms of unionism might have emerged. The tragedy is that we can never know what might have been discovered in the absence of these government-imposed blocks to entrepreneurial



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

discovery. All we know is that we are the victims of undiscovered discovery.

Finally, as Mises (1949) pointed out, all people behave entrepreneurially. Workers, no less than employers and investors, undertake to discover and correct economic error when it is in their interest to do so, and when the costs of doing so are low enough to make it worthwhile. Workers hired to replace strikers could be a major source of entrepreneurial discovery of ways and means of increasing productivity, improving the workplace environment, settling disputes, and improving labor-management relations. The job property rights granted to strikers makes hiring replacement workers more difficult than it otherwise would be. Workers are reluctant to accept jobs they know are temporary, and if they accept the jobs, they will not be motivated to be as alert to possibilities for long-run organizational improvements as they otherwise would be. Undiscovered and stifled discovery again take their toll.

## The Schlossberg-Fetter View

The SF study (1986) is an invitation to practitioners and analysts of labor law to address the issues posed by the apparent points of conflict between labor law and the development of new forms of labor-management cooperation. In their words (pp. 2–3, 12):

> This potential conflict . . . has led the U.S. Department of Labor to embark on a study of the nation's labor and related laws, and, perhaps, also collective bargaining traditions and practices that may inhibit improved labor-management relations. . . .
>
> [W]e will seek to determine whether our existing labor laws impede or, indeed, totally bar, many of the cooperative efforts the Department is currently encouraging and publicizing. . . .
>
> Clearly . . . it seems unproductive and even foolish to allow our traditional adversarial model of industrial relations to retard the spread of innovative and promising cooperative employer-employee ventures.

Lest these words be misconstrued to imply that SF are looking to the repeal of the laws that have given us our "traditional adversarial" labor relations, or even to amendments of the laws that would make them less adversarial, consider their discussion and their approach.

After briefly describing what they consider to be favorable programs of union-management cooperation at GM's Fiero plant in Pontiac, Michigan, and the NUMMI plant in Fremont, California, they come to the GM Saturn Project. They characterize the Saturn agreement between GM and the UAW as "an extraordinary commitment to cooperative ways of dealing," which is "the longest step yet taken [by management] toward full partnership with labor in every phase



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of planning and production," wherein "the employer gains maximum flexibility and employees secure new forms of compensation and employment protection" (p. 12). SF lament that this "innovative Saturn understanding and a forward-looking agreement between the building trades unions and their contractors were threatened" by charges that the agreements violated current labor law (p. 12). They imply that the work rules and compensation schemes agreed to by the UAW and GM were challenged as unfair labor practices.

In fact no such allegations were made. The National Right to Work Legal Defense Foundation charged that the prehire agreement, whereby union members were given hiring preference over nonunion workers and GM agreed to recognize the UAW as exclusive bargaining agent for all workers before letting the workers vote on the matter, violated Sections 9(a) and 8(a)3 of the NLRA. The Foundation expressed no disagreement whatsoever with the forms of cooperation adopted or with the idea of seeking cooperative solutions, but it strongly objected to coercing yet-to-be-hired workers to accept the UAW as exclusive bargaining agent and to discriminating against nonunion workers at initial hiring.

The actual nature of SF's concern is made clear by their reference to the "forward-looking agreement" between the building trades unions and contractors. That agreement was simply that only union labor would be used in the construction of the Saturn plant. Nonunion firms would not even be permitted to bid on the project. The whole of the SF essay focuses on the issue of *union*-management cooperation, not on the issue of labor-management cooperation. This comes as no surprise to readers who know that Schlossberg was formerly chief counsel for the AFL-CIO. He quite naturally looks at all labor issues through union glasses. To SF, cooperation is fine, as long as it involves an exclusive bargaining agent. Cooperation of employers with nonunion employees has no value at all.

The same perspective is apparent in SF's discussion of the relevance of *NLRB v. Yeshiva University* (444 U.S. 672 [1980]) to the issue of labor-management cooperation. In that case the Supreme Court declared that faculty members at private colleges and universities, who exercise decision authority over faculty hiring and promotion, as well as curriculum and other management-type matters, cannot be covered by the NLRA. SF explicitly align themselves with the views expressed by John T. Dunlop, former Secretary of Labor. Dunlop laments that the *Yeshiva* decision is an inappropriate "transfer to the academic or to the medical care worlds of the industrial plant model of supervision or hiring, promotion and discharge."[5]

[5]Quoted by SF (1986, p. 11).



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Now, I certainly agree that the industrial model should not be transplanted to the academy. But to me that means that the whole apparatus of monopoly compulsory bargaining set up in the NLRA has no place in colleges and universities (Baird 1981). This is not what SF and Dunlop have in mind. They want exclusive representation and mandatory bargaining in the academy. What they do not want the academy to adopt from the industrial model is the distinction between management and nonmanagement employees because that distinction, under current law, permits private college and university faculties to escape from monopoly compulsory bargaining. They fear that, in light of newly emerging forms of labor-management cooperation that blur the distinction between management and nonmanagement employees, still more employees will be able to avoid NLRA unionism.

SF evaluate the *Yeshiva* decision in these words (pp. 14–16):

> In the final analysis, the Court finds inconsistent with the NLRA's model any requirement for bargaining between the real management of an institution and the faculty. That conclusion, unfortunately, jeopardizes the desired method of operation in an ideal employee/employer participation plan, and, if extended, could potentially cast doubt upon the most innovative features of the Saturn, NUMMI, and Pontiac Fiero relationships. . . .
>
> [W]hen the NLRB finds . . . that physicians and dentists who have a say in the management of their facility cannot belong to a union, those individuals best suited to helping labor-management cooperation along are lost to the cause. . . .
>
> In the end the *Yeshiva* case is troubling because it is at war with the idea of consensus between professional employees and their administrators—and, by analogy, in other employee-employer relationships as well. That, of course, is the nub of the situation. Do our labor laws and the decisions they spawn prevent or impede problem solving and the advantages which flow from such an orientation?

Why should the "ideal employee/employer participation plan" require an exclusive bargaining agent representing employees? The nonunion collegial form of university governance is squarely based on the idea of cooperation and consensus between administrators and professionals. Faculty unionism in public universities has nearly destroyed that idea. Who says that the "individuals best suited to helping labor-management cooperation along" are those of unionist sympathies? And if they are "lost to the cause," what cause? The cause of cooperation, or the cause of unionism? Is the *Yeshiva* decision "at war with the idea of consensus between professional employees and their administrators" or at war with the extension of monopoly compulsory bargaining?



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

There follows a long discussion of what SF call "the 8(a)2 problem" related to the *Cabot Carbon* case. They recognize, as I do, that the development of employee participation in management and quality of work-life arrangements is likely to be impaired by the NLRA's strictures against company-sponsored "labor organizations," but their remedy is not to do away with those strictures. They argue (p. 16) that

> there is an obvious necessity to preserve genuine prohibitions on "company unionism" and to proscribe the use by unscrupulous employers of spurious cooperative or participation schemes conceived as means to defeat legitimate efforts at union organization. Labor-management cooperation in the context of this paper contemplates and assumes the coming together of truly independent entities in a cooperative activity perceived to be mutually beneficial.

SF simply assume that only NLRA unions can be "truly independent." Any labor-management cooperation programs in the absence of such unions are, in their view, likely to be "schemes conceived as means to defeat legitimate efforts at union organization." SF do not tell us how the law should be changed to cope with the 8(a)2 problem, they simply tell us how it ought not to be changed.

Then there is the matter of the distinction between mandatory and permissive subjects of bargaining. In *NLRB v. Wooster Division of the Borg-Warner Corp* (356 U.S. 342 [1958]) the Supreme Court held that the duty to bargain in Section 8(a)5 of the NLRA is limited to those areas specifically listed in Section 9(a)—"rates of pay, wages, hours of employment, or other conditions of employment." Employers must bargain with exclusive bargaining agents on those matters. They are mandatory subjects of bargaining. But on other matters, such as partial or complete plant closings for economic reasons, consolidations of operations, and benefits for already-retired employees, employers do not have to bargain with unions. They may do so if they choose, but they do not have to. The latter are permissive subjects of bargaining. Unions may not bargain to an impasse and strike over matters that are merely permissive subjects of collective bargaining. Decisions on permissive subjects of bargaining are considered management prerogatives.

About this distinction SF write (p. 25):

> Only by removing the artificial distinction between mandatory and permissive subjects of bargaining can both parties be assured that any issue relevant to the employment relationship will, at the behest of either side, receive full and fair discussion and consideration. . . . [S]hould company and union representatives argue about what they



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> *must* discuss, or should they be encouraged to find solutions to *whatever* perceived problems between them they can lawfully resolve?

Although SF do not say so, they strongly imply that the best remedy here is to make all subjects mandatory subjects of bargaining. If the law were changed in this way it would mean that, where there was an exclusive bargaining agent, no decisions regarding private firms could be made without the concurrence of the union. Management, representing owners, would have to get union permission even to go out of business. Effective private ownership and operation of business would cease to exist. There would be a de facto socialization of ownership and control.

That SF would consider such an outcome acceptable is not altogether surprising. In the Dunlop speech referred to above, the former Secretary of Labor states that he thinks the "needs of the country" sometimes require "tripartite consultations" involving unions, management, and government, "over such questions as shutdown of operations or plants" (Dunlop 1984, p. E-3). The fact that "tripartite consultations" were the main feature of Mussolini's organization of the Italian economy is instructive.

My solution to any problems created by the distinction between mandatory and permissive subjects of bargaining is to make bargaining on all subjects *voluntary*. If the end is to develop labor-management cooperation, the extension of coercion would seem to be a singularly inappropriate means. The essence of cooperation is voluntary exchange. Mutually beneficial arrangements will systematically emerge only when all parties are at liberty to say no and walk away.

## Conclusion

There is indeed an inherent conflict between existing labor law and innovations in labor-management relations based on cooperation rather than conflict. The NLRA was fashioned by people who thought of labor-management relations as inherently adversarial. In fact it is only compulsory union-management relations that are unavoidably adversarial, and that adversity is largely responsible for the inability of many basic American industries effectively to compete. The obvious remedy is to repeal the NLRA and return to the form of unionism President Franklin Roosevelt endorsed in his March 1934 Automobile Settlement—voluntary unionism based on proportional representation and the right of workers to choose, on an individual basis rather than by majority vote, whether to be represented by a union



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and, if so, whether the union is independent or employer-sponsored (Baird 1984, pp. 39–42).

SF, however, have a different agenda. Under the guise of promoting labor-management cooperation they apparently want to resurrect the failed 1978 labor law reform movement (p. 30):

> These difficult issues and the manner in which they are handled will be crucial in setting the future tone of U.S. labor relations. . . . It has been clear since the union-backed labor reform effort failed in 1978 that, in the words of Professor Dunlop, "redesign must come from discussion and consensus among labor [unions] and management, principals not lawyers, to be followed by public and legislative review."

In other words, in 1978 the politicians failed to give us what we want so now we must take a more roundabout route to our goals.

The bottom line of the 1978 reform effort was to make it easier for unions to become certified exclusive bargaining agents. SF develop this theme by favorably referring to Paul Weiler's proposed changes in labor law. Weiler (1983), deeply concerned about the recent union failure rate in representation elections, proposed, among other things, that such elections be held immediately after unions apply to the NLRB for them, before employers have a chance to campaign against union representation. But surely it makes little sense to pursue competitive advantages by expanding the very form of unionism that is largely responsible for our present competitive disadvantages.

## References

Baird, Charles W. "Some Philosophical and Economic Arguments in Support of Union-Free Colleges and Universities." *Government Union Review* 2 (Spring 1981): 22–30.

Baird, Charles W. *Opportunity or Privilege: Labor Legislation in America.* Bowling Green, Ohio: Social Philosophy and Policy Center, Bowling Green State University, 1984.

Baird, Charles W. "Comparable Worth: The Labor Theory of Value and Worse." *Government Union Review* 6 (Winter 1985a): 1–29.

Baird, Charles W. "Labor Law and Entrepreneurial Discovery." Ludwig von Mises Institute, *Austrian Economics Newsletter* (Fall 1985b): 1–5.

Baird, Charles W. "Entrepreneurial Discovery in the Regulated Labor Market." In *The Age of the Entrepreneur: Private Initiative for Economic Development.* Edited by Michael M. Kurth. Bryn Mawr, Pa.: Intercollegiate Studies Institute, forthcoming (1987).

Dunlop, John T. An Address on "Legal Framework of Industrial Relations in the United States." Bureau of National Affairs, *Daily Labor Report* (7 October 1985): E-1–E-5.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hayek, F. A. "The Use of Knowledge in Society" [1945]. In *Individualism and Economic Order*, pp. 77–91. Edited by F. A. Hayek. Chicago: University of Chicago Press, 1948.

Hayek, F. A. "Competition as a Discovery Procedure "[1968]. In *New Studies in Philosophy, Politics, Economics, and the History of Ideas*, pp. 179–90. Edited by F. A. Hayek. Chicago: University of Chicago Press, 1978.

Kirzner, Israel M. *Competition and Entrepreneurship*. Chicago: University of Chicago Press, 1973.

Kirzner, Israel M. *Discovery and the Capitalist Process*. Chicago: University of Chicago Press, 1985.

Mises, Ludwig von. *Human Action*. New Haven: Yale University Press, 1949.

Nelson, Daniel. "The Company Union Movement, 1900–1937: A Reexamination." *Business History Review* 56 (Autumn 1982): 335–57.

Reynolds, Morgan O. *Power and Privilege*. New York: Universe Books, 1984.

Reynolds, Morgan O. "Unions and Jobs: the U.S. Auto Industry." *Journal of Labor Research* 7 (Spring 1986): 103–26.

Schlossberg, Stephen I., and Fetter, Steven M. *U.S. Labor Law and the Future of Labor-Management Cooperation*. Washington, D.C.: U.S. Department of Labor, BLMR 104, 1986.

Weiler, Paul C. "Promises to Keep: Securing Workers' Rights to Self-Organization under the NLRA." *Harvard Law Review* 96 (June 1983): 1769–1827.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# BOOK REVIEWS

**Perspectives on Safe and Sound Banking: Past, Present and Future**
George J. Benston, Robert A. Eisenbeis, Paul M. Horvitz,
Edward J. Kane, and George G. Kaufman
Cambridge: Mass.: MIT Press for the American Bankers Association, 1986, 358 pp.

This is an important book. Commissioned by the American Bankers Association (ABA), George Benston, Robert Eisenbeis, Paul Horvitz, Edward Kane, and George Kaufman conduct a stimulating examination of the industry. Their work is much more than a set of loosely related essays brought together by an editor, however; it is a fully integrated book. The drafts of each chapter were circulated to the entire team for comments and reconciliation of views. Their perspectives provide a sound, frequently innovative, and persuasively argued analysis of bank stability.

The authors cover a wide range of topics: the lack of diversification in bank portfolios (Chapter 1), excessive risk-taking associated with mispriced deposit insurance (Chapter 3), the role of the lender of last resort (Chapter 5), the expansion of bank powers (Chapter 6), the value and probable efficacy of various kinds of market discipline as a tool for controlling bank risk (Chapter 7), the potential advantages of introducing risk-based deposit insurance premiums (Chapter 9), the high cost and misdirection of the standard bank field examinations (Chapter 10), and past and present proposals for reforming the bank regulatory system (Chapter 11). In addition to describing the current state of the industry and the nature of the debate over banking regulation, there are many important new contributions made to the discussion of the proper role of government regulation.

In Chapter 2, "The Consequences of Bank Failure," for example, George Kaufman provides some bracing analysis against sloppily based fears of bank runs. Kaufman shows that except during the years of the Great Depression, bank failures have been less common and have imposed smaller losses on depositors than is generally supposed.

*Cato Journal*, Vol. 6, No. 3 (Winter 1987). Copyright © Cato Institute. All rights reserved.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Kaufman is somewhat ambivalent, however, about the relationship between the extent of branching and concentration in the banking system and the system's stability. While he follows W. F. Spahr in attributing Canada's escape from bank failures during the 1930s to its nationwide branching system (p. 68), Kaufman also ascribes the fears of banking contagion expressed by 19th century British writers Henry Thornton and Walter Bagehot to the highly concentrated nature of British banking.[1] He argues that the large "number of banks in the United States, and the unit banking structure" insulate the United States from a contagious "public loss of confidence in the banking system" (p. 71). It is unclear, therefore, whether Kaufman believes branching and concentration promote or impede bank safety.

In Chapter 4, Horvitz addresses "Alternative Ways to Resolve Insolvencies." He describes the many methods for failure resolution used by the Federal Deposit Insurance Corporation (FDIC) in the past. The authors adopt a recommendation proposed elsewhere by Kaufman—that of a modified trusteeship program. The insurer would act promptly when a bank reaches insolvency, writing down uninsured deposits overnight by a "haircut" that would represent the agency's estimated maximum loss. Next day the bank would reopen, administered by the insurers pending a more permanent resolution. Stockholders and management would be more severely penalized than at present in order to discourage risk taking. The study might have, but unfortunately did not, recommend regularly giving taxpayers equity shares in rescued institutions. No doubt the authors want to avoid nationalization, but this can, and should, be precluded by other means.

In Chapter 5, Edward Kane and George Kaufman discuss the role of "The Lender of Last Resort." Surprisingly these authors state that a "threat of simultaneous insolvencies at a large number of sizable institutions" which "might cause systemwide difficulties" could justify lender of last resort (LLR) assistance to insolvent institutions (p. 116).

Widespread simultaneous insolvencies would be unlikely if monetary policy were always conducted correctly. Unfortunately, that is not the case. The impact of recent disinflation, for example, on the agricultural, energy, old manufacturing, and international debt portfolios of banks is portentous. The challenge facing the savings and loan (S&L) industry for the past five years has been widespread insolvencies. Are the authors, in fact, recommending prolonged and widespread LLR assistance to banks and S&Ls facing current difficulties?

In Chapter 8 on "Market-Value Reporting," Kane makes a vigorous appeal for using market value accounting (MVA) measures instead of book values. He argues the change would improve management and benefit stockholders, the public, and the regulators. Kane acknowledges the practical difficulties—the lack of secondary market data—but he recommends an industry self-regulatory body to help resolve the problems. His enthusiasm evidently is not shared by all team members, however, as the team's final recommenda-

[1]The British system was much less concentrated in the 19th century than it is now, but still appeared compact in comparison with the U.S. system.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tion (in the appendix) is to permit, but not require, MVA both for internal use and public disclosure. MVA is recommended for insurance purposes, representing a significant departure from present practice.

Rather like an old soldier, the book fades away in Chapter 11. But it is surprisingly resurrected in a vigorous appendix that provides the authors' recommendations for "Improving Efficiency, Safety and Soundness of the Banking System."

Closer integration of the recommendations into the book would have strengthened it substantially. As it is, some of the recommendations come as a surprise to the reader. For example, the recommendation concerning risk-based capital requirements involves a subject virtually omitted from the book.

A second odd thing about the book is its focus on banking. Can the stability of the whole financial system be equated with that of banks? Since the 1980 and 1982 deregulation acts, banks have become less "special," so that a broader discussion of thrift issues, at least, would have been appropriate. The concentration on banking is acknowledged in the book's title, however, and it is understandable given the ABA's role.

There is at least one mistaken policy recommendation caused in part by the failure to consider the problems facing S&Ls, however. The authors recommend that the insuring agency, instead of the Federal Reserve, should lend to institutions experiencing liquidity problems (p. 313). The intention is to prevent "insolvent institutions from being subsidized and imposing losses on the insurance funds." But the Federal Savings and Loan Insurance Corporation (FSLIC) has not followed its own rules and closed failed thrifts. Instead the Federal Home Loan Banks have lent heavily to insolvent S&Ls, allowing them to continue incurring losses that add to the future burdens of the FSLIC, the industry, and the taxpayer.

The reader also may be unconvinced by the assertion that "a system of risk-related premiums needs not be actuarially perfect in order to be a useful constraint on risk-taking" (p. xix). Are the authors of this book fully convinced that even inappropriate premiums could make the future situation no worse than the present? Or do they believe the regulators are unlikely to make mistakes in setting insurance premiums and that appropriate premiums can be costlessly determined? Or do they think that *ex post* settlement will offset any undesirable incentives?

Readers who favor deregulation may be disturbed by the basic premise of the book that the central bank, the systems of deposit insurance, examination, supervision, and regulation are all necessary. Indeed, the authors do not debate this issue; they merely state at the outset that "the efficient functioning of the financial system benefits from appropriate government intervention" (p. xv). Nevetheless, the authors take a relatively free-market position: "[M]uch of this report considers the means of increasing the effectiveness and minimizing the extent of such intervention, and of harnessing market forces and incentives to maintain bank stability" (p. xv).

There are some notable omissions from the book. No mention is made of international banking and the developing nations' debt crises. Nor is there



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

any real discussion of the rapid growth and increasing risk of off-balance sheet liabilities. Finally, a brief discussion of the securitization of bank portfolios, the bundling and selling of bank loans in secondary markets, and their corresponding impact on liquidity and the need for market value accounting would have strengthened the book.

In summary, this is an authoritative, vigorous, albeit not exhaustive, study that results in a wealth of recommendations whose implementation could greatly ease the present situation and prevent it from deteriorating further. Students, practitioners, and scholars will all benefit by reading it.

Gillian Garcia and Kim Staking
U.S. General Accounting Office

**Superfairness**
William J. Baumol
Cambridge, Mass.: MIT Press, 1986, 266 pp.

We have long been familiar with fairness as a normative concept. As such, its application was capricious and its implications acceptable only to the extent one accepted the analyst's definition. In this book, William Baumol extends the works of Duncan Foley, E. A. Pazner and David Schmeidler, and Hal R. Varian to provide a positive, analytically useful theory of fairness.[1]

The book, essentially, has two sections; one presenting the theory and the other presenting applications. Chapters 1 through 3 contain the positive theory of superfairness. A distribution of commodities is said to be "superfair" if no participant prefers another's share to his own, that is, if each participant (strictly) prefers his share to that of all other participants. For purposes of application, these definitions are extended. "Incremental superfairness" implies a change in distribution in which no participant prefers another's increments to his own. "Partial superfairness" and "partial incremental superfairness" apply the above definitions to proper subsets of all commodities.

Chapters 4 through 12 contain the applications of the theory of fairness. In Chapter 4, "On Rationing of Scarce Commodities," Baumol draws from a previous article to present an extensive and appealing application of the fairness criterion. Other applications include compensation in the face of negative externalities (Chapter 5), pricing in multiproduct firms (Chapters 6 and 7), peak and off-peak pricing (Chapter 8), taxation and subsidization (Chapter 9), and arbitration (Chapter 12). Baumol also suggests the application of fairness theory to divorce settlements (Chapter 11). Chapter 10 deals with the prevalence of economic illusion.

[1]See Duncan Foley, "Resource Allocation and the Public Sector," *Yale Economic Essays* 7 (Spring 1967): 45–98; E. A. Pazner and David Schmeidler, "A Difficulty in the Concept of Fairness," *Review of Economic Studies* 41 (July 1974): 441–43; and Hal R. Varian, "Equity, Envy and Efficiency," *Journal of Economic Theory* 9 (September 1974): 63–91.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The more tractable definitions of incremental and partial fairness, which are utilized in the applications, are somewhat tainted. Baumol demonstrates that "a move that is both incrementally superfair and a Pareto optimal improvement from an initial position that is superfair can yield a new position that is unfair" (p. 57). Similarly, "a distribution that satisfies a partial fairness . . . criterion is not necessarily fair" (p. 64). Thus, the definitions of fairness that are so philosophically appealing are not necessarily present in the analysis of fairness. For example, price discrimination may prove to be incrementally fair while the resulting distribution is unfair and vice versa.

Aside from the impossibility of inferring fairness from partial or incremental fairness, the concept of fairness in the distribution of commodities has been criticized for ignoring the *process* by which the distribution is generated. In a comment on Baumol's earlier work, Randall Holcombe has noted, "If one person builds two chairs in a day when an equally able person spends the day watching television, it would be fair for the first person to keep the two chairs; yet if their bundles were equally desirable at the beginning of the day, the fair solution according to Baumol would be for each person to own one of the chairs."[2] Holcombe argues for a fair process, not necessarily a "fair" outcome.

Baumol appears to respond to Holcombe's criticism in *Superfairness* by introducing the concept of "contribution fairness" in which no participant prefers another's consumption and labor contribution. Implicitly, when production, saving, or uncertainty is involved, the relevant point for the applicant of fairness theory is *ex ante*. The loser prefers the winner's poker hand. The preponderant consumer prefers the preponderant saver's retirement plan. Yet, presumably, each preferred the process that led to their position. While Baumol pays lip service to the fairness of the process through his introduction of contribution fairness, none of his applications is presented *ex ante* and the concept of contribution fairness is never applied.

For example, in his discussion of the fairness of various rationing schemes, producers of the scarce commodity are not considered. Nor are speculators, who might have stockpiled the commodity in anticipation of rising prices or shortages, considered. The analytical reference point is *after* the imposition of a ceiling price. Similarly, following a demonstration that the use of progressive taxation and transfers to promote absolute income equality would lead to zero output, Baumol considers varying wages across individuals in such a way that absolute income equality is achieved with positive output. The purpose, ostensibly, is to generate a fair distribution that does not seriously reduce output. Once again, the analytical reference is clearly *ex post* and the process is ignored. While no one prefers another's take-home pay *ex post*, many would prefer another's wage rate *ex ante*.

Lest I appear negative with regard to *Superfairness*, let me point out that my only criticism hinges on the *ex post, ex ante* positioning of the application of the theory. If care is taken, fairness analysis has a bright future. This is

[2]Randall G. Holcombe, "Applied Fairness Theory: Comment," *American Economic Review* 73 (December 1983): 1153.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

precisely the point Baumol makes about the application of the partial and incremental fairness criterion. I anticipate that in the future, contribution fairness will replace distributional fairness as the dominant measure of fairness.

Philip Porter
University of South Florida



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

# Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify)_____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name_____
Title_____
Institution/Company_____
Department_____
Address_____
City_____ State_____ Zip_____
Phone (_____)_____

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106



# Southern Economic Journal

*Joint Publication of the Southern Economic Association and the University of North Carolina*

The *Southern Economic Journal* is published four times a year, in January, April, July, and October, at Chapel Hill, N.C. 27514. The subscription price is $25.00 per year for members; $40.00 for nonmembers, domestic, and $46.00, foreign. Single copies: current issue $10.00 domestic, $11.50 foreign; back issues $11.00 domestic, $12.50 foreign. Circulation 4000. Address communications to:

Vincent J. Tarascio, Managing Editor
Southern Economic Journal
300 Hanes Hall, 019A
Chapel Hill, N.C. 27514

A $15.00 submission fee for SEA members and $25.00 for nonmembers must accompany all manuscripts except book reviews. Membership dues in the Association are: annual $25.00, contributing $42.50, life $325.00, institutions $325.00, student $12.50, retired $12.50. Membership dues include payment for a year's subscription to the *Southern Economic Journal*. Applications for membership should be addressed to Joseph Jadlow, College of Business Administration, Oklahoma State University, Stillwater, Oklahoma 74078.



LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

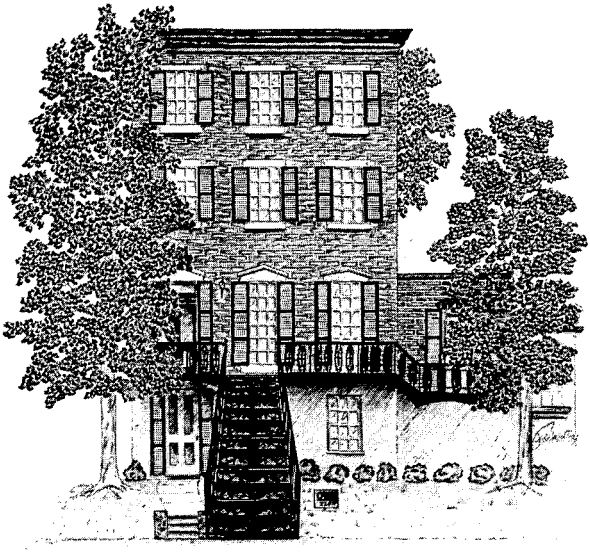
# The Cato Institute

Founded in 1977, the Cato Institute is a nonpartisan public policy research foundation headquartered in Washington, D.C. The Institute is named for *Cato's Letters,* libertarian pamphlets that helped lay the philosophical foundation for the American Revolution.

The Cato Institute seeks to broaden the parameters of public policy debate to allow consideration of more options that are consistent with the traditional American principles of limited government, individual liberty, and peace. Toward that goal, the Institute strives to achieve a greater involvement of the intelligent, concerned lay public in questions of policy and the proper role of government.

The Cato Institute undertakes an extensive publications program dealing with the complete spectrum of policy issues. Books, monographs, and shorter studies are commissioned to examine the federal budget, Social Security, monetary policy, natural resource policy, military spending, regulation, NATO, international trade, and many other issues. Major policy conferences are held throughout the year, from which papers are published thrice yearly in the *Cato Journal.*

In order to maintain an independent posture, the Cato Institute accepts no government funding. Contributions are received from foundations, corporations, and individuals, and from the sale of publications. The Institute is a nonprofit, tax-exempt, educational foundation under Section 501(c)3 of the Internal Revenue Code.



The Watterston House headquarters of the Cato Institute

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

$6.00

# PROPERTY, JUSTICE, AND GOVERNMENT

Government is instituted to protect property of every sort; as well that which lies in the various rights of individuals, as that which the term particularly expresses. This being the end of government, that alone is a *just* government, which *impartially* secures to every man, whatever is his *own*. . . .

That is not a just government, nor is property secure under it, where the property which a man has in his personal safety and personal liberty, is violated by arbitrary seizures of one class of citizens for the service of the rest. . . .

That is not a just government, nor is property secure under it, where arbitrary restrictions, exemptions, and monopolies deny to part of its citizens that free use of their faculties, and free choice of their occupations, which not only constitute their property in the general sense of the word; but are the means of acquiring property strictly so called. . . .

If there be a government, then, which prides itself in maintaining the inviolability of property; which provides that none shall be taken *directly* even for public use without indemnification to the owner, and yet *directly* violates the property which individuals have in their opinions, their religion, their persons, and their faculties; nay more, which *indirectly* violates their property in their actual possessions, in the labor that acquires their daily subsistence, and in the hallowed remnant of time which ought to relieve their fatigues and soothe their cares, the influence [inference] will have been anticipated, that such a government is not a pattern for the United States.

If the United States mean to obtain or deserve the full praise due to wise and just governments, they will equally respect the rights of property, and the property in rights: they will rival the government that most sacredly guards the former; and by repelling its example in violating the latter, will make themselves a pattern to that and all other governments.

**—James Madison**

"Property." *National Gazette*
1, no. 44 (29 March 1792):
174–75.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED